4

RADC-TR-89-93
In-House Report
August 1989

KNOWLEDGE BASED TEXT GENERATION

Mark T. Maybury, 1Lt, USAF

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, NY 13441-5700

This report has been reviewed by the RADC Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RADC-TR-89-93 has been reviewed and is approved for publication.

APPROVED:

SAMUEL A. DINITTO, Jr.
Chief, Command & Control Software Technology Division
Directorate of Command and Control

APPROVED:

RAYMOND P. URTZ, Jr.
Technical Director
Directorate of Command and Control

FOR THE COMMANDER:

IGOR G. PLONISCH
Directorate of Plans and Programs

DESTRUCTION NOTICE - For classified documents, follow the procedures in DOD 5200.22-M, Industrial Security Manual, or DOD 5200.1-R, Information Security Program Regulation. For unclassified limited documents, destroy by any method that will prevent disclosure of contents or reconstruction of the document.

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC (COES) Griffiss AFB, NY 13441-5700. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS N/A
2a. SECURITY CLASSIFICATION AUTHORITY N/A 2b. DECLASSIFICATION / DOWNGRADING SCHEDULE N/A	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
4. PERFORMING ORGANIZATION REPORT NUMBER(S) RADC-TR-89-93	5. MONITORING ORGANIZATION REPORT NUMBER(S) N/A

6a. NAME OF PERFORMING ORGANIZATION Rome Air Development Center	6b. OFFICE SYMBOL *(If applicable)* COES	7a. NAME OF MONITORING ORGANIZATION Rome Air Development Center (COES)
6c. ADDRESS *(City, State, and ZIP Code)* Griffiss AFB NY 13341-5700		7b. ADDRESS *(City, State, and ZIP Code)* Griffiss AFB NY 13441-5700
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Rome Air Development Center	8b. OFFICE SYMBOL *(If applicable)* COES	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N/A
8c. ADDRESS *(City, State, and ZIP Code)* Griffiss AFB NY 13441-5700		10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO
62702F	5581	27	TK

11. TITLE *(Include Security Classification)*
KNOWLEDGE BASED TEXT GENERATION

12. PERSONAL AUTHOR(S)
Mark T. Maybury, 1Lt, USAF

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT *(Year, Month, Day)*	15. PAGE COUNT
Final	FROM Oct 87 TO Oct 88	August 1989	110

16. SUPPLEMENTARY NOTATION
N/A

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
12	05	

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Natural Language Generation Text Planning
Knowledge Based Systems GENNY
Computational Linguistics Rhetorical Schema

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

This report develops a theory of text generation and describes an implemented computational model of this theory. The theory attempts both domain independency at the knowledge level and language independency at the linguistic level by drawing and expanding upon previous work in discourse schema and grammatical relations, respectively. The implemented system, GENNY, generates coherent texts by employing discourse strategies (which occur in human produced text) in parallel with pragmatic constraints (e.g., focus and context).

The report begins with an introduction and summary of the research performed. This is followed by a survey of the text generation literature which places GENNY in the context of past language production research. Next, the motivation for the theoretical position adopted is discussed followed by detail of the theory on a knowledge, pragmatic, semantic, relational, and syntactic level, with illustration of the practical implementation throughout. Results of GENNY's test production from two frame (Cont'd)

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL MARK T. MAYBURY, 1Lt, USAF	22b. TELEPHONE *(Include Area Code)* (315) 330-7794	22c. OFFICE SYMBOL RADC(COES)

DD Form 1473, JUN 86 Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

19. (Cont'd). knowledge bases (neuropsychology and photography) are then presented together with preliminary interlingual test results (English and Italian). GENNY is evaluated with respect to state-of-the-art generators and is shown to be equivalent and, in some respects, superior, in competence and performance. In conclusion, the contributions and limitations of the system are discussed and areas for further development are suggested.

Contents

Preface

This work is a revised version of a dissertation submitted in partial fulfillment of the requirements for the degree of Master of Philosophy in *Computer Speech and Language Processing* at Cambridge University, England.

Special thanks to my supervisor, Professor Karen Sparck Jones, for perspicacious discussions and encouragement. Also, to Professor Frank Fallside for providing superb research facilities (and nice chairs!). To Dr. Steve Pullman and Dr. Steve Young for guidance. To Neil Russell, Dr. Nick Youd and the other Alvey Project members, and the Engineering lab Ph. D. students for emotional support and technical advice. And to the M. Phil. clan for late night hacking, early morning theorization, and light tea-room conversation.

This work was funded by an *Ambassador of Goodwill* postgraduate scholarship from the Rotary Foundation, International. Thanks to the United States Air Force Institute of Technology for financial and administrative support as a student officer and to my program manager, Lt. Col. Parsons, AFIT/CI.

Keywords: natural language generation, knowledge representation, expert systems, focus, discourse, pragmatics, semantics, syntax.

Chapter 1

INTRODUCTION

I have laboured to refine our language to grammatical purity,and to clear it from colloquial barbarisms, licentious idioms, and irregular combinations.
Samuel Johnson, 1752

The central aim of natural language generation (NLG) is to investigate the knowledge and processes -- both linguistic and extra-linguistic -- that speakers and writers employ in order to communicate to their intended audience. Production, therefore, encompasses issues of deciding what is pertinent as well as determining how to organize and present information effectively. Speakers and writers must also select proper words and form appropriate sentence structures. These issues are manifest in the questions:

- *What* should we speak about?
- *When* should we speak about it?
- *How* should we speak about it?

In this work, a linguistic approach and computational system (GENNY) are presented which offer a framework from which the generation process -- what, when and how -- can be investigated. In particular, this work addresses the issues of pertinency, coherency and grammaticalness and demonstrates algorithms and mechanisms for achieving these.
This discussion, therefore, encompasses not only traditional issues of syntax and semantics, but equally current problems in pragmatics and discourse theory (e.g. supra-sentential connectivity of

text). GENNY demonstrates how these higher level constraints can effect the low-level realization of language in a well-motivated manner.

1.1 Summary/Example

GENNY was built to answer general questions about both the permanent structure of a knowledge base as well as the results of an individual run of an expert system in neuropsychology. Three type of wh interrogatives were addressed: queries for definitions (What is an X?), requests for explanations (Why did you obtain the result Y?), and requests for comparisons (What is the difference between X and Y?). For example, asked to define a brain, GENNY responds:

```
A brain is a region for understanding located in the human skull.
It has a relative importance value of ten.[1]
It contains two regions: the left-hemisphere region and the
right-hemisphere region.
The left-hemisphere has a relative importance value of ten.
The right-hemisphere has a relative importance value of ten.
The right-hemisphere region, for example, has the gestalt-
understanding function located in the right brain.
```

After loading a new knowledge base[2] and dictionary on photography, we ask GENNY to explain why the expert system diagnosed a camera aperture fault. She responds:

```
The aperture component is damaged because the light-pictures
observation and the dark-pictures observation indicate damage.
The light-pictures observation has a likelihood value of six.
The dark-pictures observation has a likelihood value of eight.
```

Figure 1.1 illustrates the knowledge and processes engaged during generation. Language input is simulated by a menu which offers the user a choice of a discourse goal (define, explain, or compare) and then asks for a specific frame in the knowledge base which serves as the discourse topic.[3] GENNY first formulates a discourse plan based on the given discourse goal. Next a relevant pool of rhetorical propositions is generated using the provided discourse topic. GENNY

[1] *Relative importance value* is the expert system representation of the significance of a piece of knowledge at some node in a generalization hierarchy with respect to its siblings.
[2] The knowledge base is the actual output from an expert system run.
[3] Actual interpretation of the discourse goal(s) and topic(s) from natural language involves non-trivial issues, but was beyond the scope of this project.

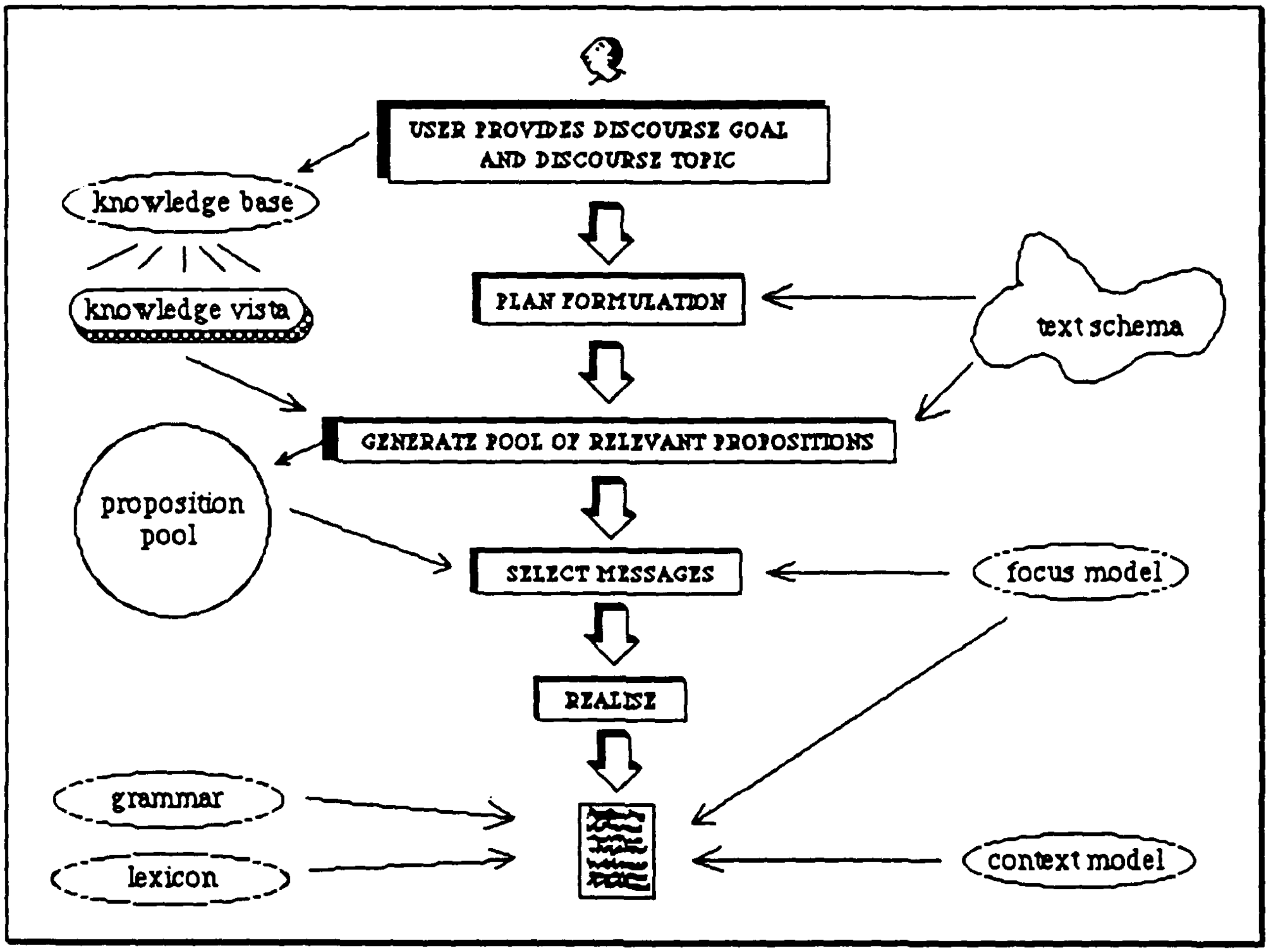

Figure 1.1 System Components and Flow of Control

then instantiates the plan -- a model of common strategies of text organization -- by choosing from among pertinent messages on the basis of pragmatic constraints of attentional focus. The subsequent list of rhetorical propositions (sequenced and connected via their linguistic role in the discourse) are translated using case semantics into a relational representation (i.e. subject, object) and then realized by a feature-based unification grammar and dictionary. Surface choice is guided by knowledge of focus (past, current, and future) and discourse context (given/new). Subsequent chapters illustrate the system components in greater detail (section 3.3 contains a theoretical overview).

1.2 Motivations

NLG is a fruitful field of endeavor because of a recent pragmatic[1] need for generation capabilities as well as the theoretical insight it offers by examining language from a non-traditional perspective: the producer. From a computational viewpoint, computer systems are increasingly dependent upon flexible natural language interfaces which accurately reflect the state of the underlying representation. This need is particularly acute in applications requiring explanatory capabilities such as expert systems or complex data bases [Malhotra, 1975].

A specific expert system, developed previously by the author [Maybury, 1986], provided direct impetus for a generation front-end. The central requirement was for a NLG system which could both educate a user about the contents of the knowledge base as well as communicate the reasoning behind a particular diagnosis. This is discussed more thoroughly in subsequent sections.

On the other hand, the theoretical aspects of how speakers identify, package and present information, involve equally non-trivial issues. The often ill-motivated linguistic components of current NLG systems suggest a need to attempt a more universal framework for production. This inadequacy manifests itself in a rough and commonly "hard-wired" transition from the planning to the realization stages. I propose to utilize relational grammar [Perlmutter, 1979, 1980, 1984], to bridge the gap between surface syntax and deep case semantics.

Moreover, it is the author's belief that attempts to develop a more coherent framework for production should offer insight into the interpretation process. While it would be oversimplifying to suggest that generation is isomorphic to interpretation, it is certainly arguable from a cognitive efficiency perspective that humans exploit nonredundant linguistic knowledge structures [Golden, 1985].[2] In this spirit, knowledge formalisms representing linguistic competence[3] (e.g. grammar

[1]The term pragmatic is used throughout this dissertation in reference to two distinct ideas depending upon context. Here it is used to mean practical or empirical whereas elsewhere it is also used to refer to the level of language which describes such phenomena as intention, belief, focus, etc. See Chapter 6 for a more detailed discussion.
[2]It may even be the case that some procedural components are shared.

and dictionary) used previously for interpretive purposes [Maybury, 1987] are here exploited for generative tasks. Hence the grammar and dictionary formalism in GENNY can be considered bi-directional. The bi-directionality of the higher level text schema remains to be investigated.

1.3 Goals

The central aim of the project was two-fold. The first goal was to develop a consistent theory of the generation process. This involved several major subtasks including: development of a domain-independent model of discourse structure based on analysis of natural texts; the identification and formulation of a (limited) set of pragmatic constraints on the generation process; and an attempt at a language-independent linguistic representation.

The second goal was to implement a computational model of the text generation process defined above to test these ideas concretely. Again, several main subtasks were involved including: analysis of natural texts and extraction of text schemata and their corresponding rhetorical predicates; design of a system motivated by the desire for domain and language independence, semantic connection of the generation system to the knowledge base (KB) formalism; implementation of algorithms constituting focus of attention; development of a unification grammar with features; coding of morphological and orthographic synthesis routines; and building of a lexical access system and a domain dictionary.

All of these initial theoretical and practical goals were met. Furthermore, an additional domain of discourse (photography) was investigated to illustrate GENNY's domain independence. Future goals of NLG research, particularly in the difficult areas of pragmatics and user modelling, were indicated.

[3]This can be contrasted with the ability to generate linguistic forms: performance.

1.4 Dissertation Organization

The typical dissertation is organized along the lines of a theoretical discussion first, followed by a description of the system implementation. In contrast, this work develops both theory and implementation in tandem. This maximizes the connection between the linguistic principles investigated and their realization in GENNY. Each section commences with a discussion of theory and background work, followed by a description of how these issues were addressed in GENNY. But in order to enhance readability, an overview of the linguistic approach (and thus dissertation organization) is presented immediately following the discussion of current research in NLG in the next chapter.

Chapter 2

TEXT GENERATION LITERATURE

Not everything is unsayable in words, only the living truth.
Ionesco

2.1 Introduction

This chapter places GENNY in the context of past and current attempts at NLG. First, early approaches to generate are reviewed. Following this is a detail of recent research which has focused on the central questions of generation: *how*, *what*, and *when* to utter. The chapter concludes by introducing the more subtle question of *why* we say something and suggests what we should do now.

2.2 Initial strategies

Initial attempts to generate language centered around single utterances in isolated context. At first messages were typed in, providing canned text as good as the human could compose. Not only does this lack flexibility, but the implementor must anticipate every necessary message and situation. This will be feasible only in the most trivial of applications. More crucially, if the underlying system is altered and the canned text remains unchanged, the actual performance of the system can be far from that which the system's messages suggest. Programmers tend to compensate for this by writing general and, oftentimes, misleading messages [Bossie and Mani, 1986].

Terry Winograd [1972] achieved a significant improvement upon canned text in his blocks world system (SHRDLU) by employing the *code conversion* technique. As the phrase implies, each entity in the underlying knowledge representation is associated with a surface text expression. These associations are manipulated with clever heuristics to map the knowledge representation onto English text. A similar direct translation of the underlying formal representation was used by Simmons and Slocum [1972 from McKeown, 1985], who grew sentences from verb case semantic networks using ATN grammars [Augmented Transition Networks from Woods, 1970].

Goldman [1975] pioneered lexical selection procedures in BABLE, the generation component of MARGIE, a system that answered questions about, made inference from, and paraphrased conceptual dependency (CD) networks [Schank, 1975]. Although he also used ATN's to generate syntactic structures, he developed discrimination net mechanisms for lexical choice. While Goldman did not linguistically justify his paraphrase choices or demonstrate contextual influence in multi-sentential output, his dictionary formulation influenced many subsequent generation systems.

These initial approaches solved some of the consistency problems of canned text since output is a product of the knowledge base.[1] Nevertheless, complex messages which require interaction of several entities in the knowledge base together with application-motivated heuristics can lead to confusing if not misleading text [Bossie and Mani, 1986].

In fact attempts at multi-utterance generation revealed the requirement for a distinct representation of linguistic knowledge. This became clear to Meehan [1979], who generated stories of goals and frustrations in his system TALE-SPIN, as well as Swartout [1981], who produced explanations from a medical consultation system. Both found that underlying knowledge formalisms were ill-suited for linguistic tasks, particularly when translating long chains of inference [c.f. McDonald, 1983]. In her system, XSEL, which helps the user produce purchase

[1] In fact, this approach is used in commercial systems (XPLAIN, EMYCIN).

orders for computer systems, Kukich [1986] suggested that messages should be generated independent from the underlying knowledge and inference mechanisms.

These representational issues were stimulated by a need for a deeper linguistic representation to deal with longer texts and the problems they entail. A computational model for text generation must incorporate mechanisms sophisticated enough to manipulate both linguistic and general knowledge to resolve the issues of *how to say something*, *when to say it*, and *what to say*..

2.3 How to Say Something

If longer texts are to be treated properly, their constituents -- *rhetorical predicates* -- must be realized in a well motivated fashion. These (ideally) domain independent messages must be mapped onto surface form with the aid of grammars, lexicons, and perhaps user models.

McDonald's [1981ab] MUMBLE generator investigated message formalisms in a variety of knowledge representations, including predicate calculus, FRL (Frame-oriented Representation Language) [Goldstein and Roberts,1977 in McDonald, 1981b] and KL-ONE, which consists of highly-structured semantic networks [BBN,1978]. This "input-driven" generator is sensitive to the previous discourse, previous decisions, as well as a user model of audience knowledge.

MUMBLE transforms this message using "two cascaded transducers folded together under the command of a single, data-directed controller" [McDonald, 1981, p. 21] (see figure 2.1). The first transducer or interpreter expands the input message into a tree which represents the surface structure. Then the controller traverses the tree depth-first[1] and uses the dictionary to replace message tokens with structure and lexical items. At the same time, the grammar is consulted to choose appropriate syntax structure. GENNY follows McDonald's philosophy of message-driven

[1]McDonald invests a significant effort into the psycholinguistic plausibility of his computational model of spoken, not written, text. Thus the decision-making process proceeds in a left to right manner. This aids efficiency significantly.

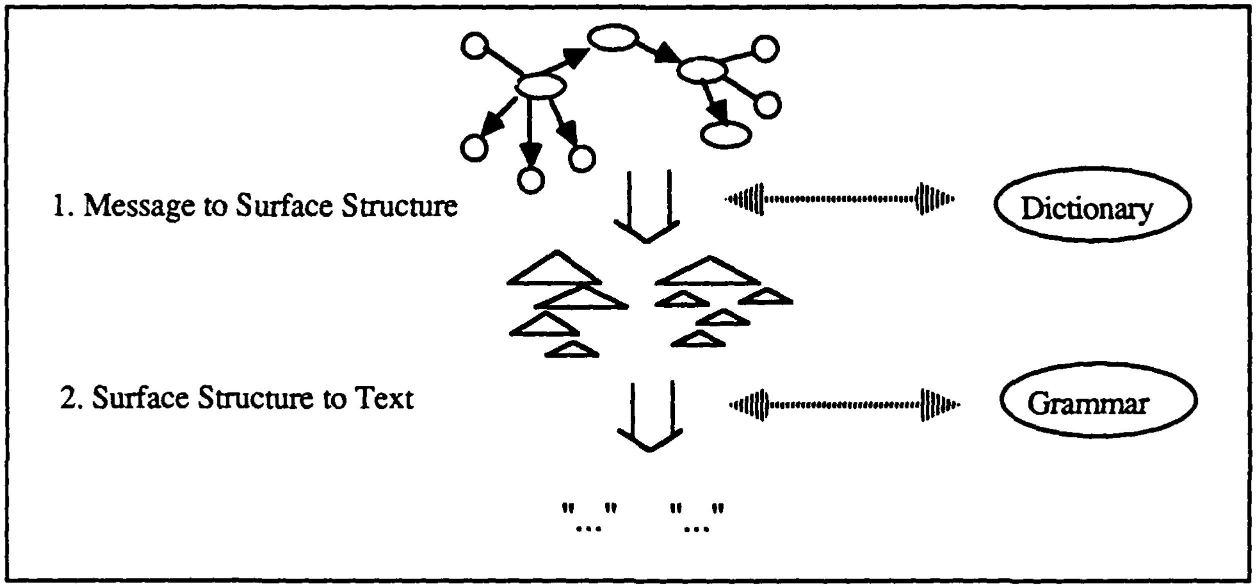

Figure 2.1 Linguistic Component [from McDonald, 1981, p. 21].

generation and syntactic independence but allows for pragmatic knowledge (e.g. focus and context) to affect surface form.

2.3.1 Grammars

Related work has focused on the development of better systemic grammars [Halliday, 1976ab]. Like ATN's,[1] systemic grammars attempt to model a system of choices, encoding grammatical aspects such as number and mood. They perform the role of GENNY's syntactic specialists. Unlike phrase structure grammars, systems are not sequentially accessed, being activated only when required. This lends efficiency and clarity.[2]

Under this formalism, [Davey, 1979] examined commentary on a game of tic-tac-toe. Davey's system had underlying concepts such as "counter-attack" and "foiled-threat" and was able to select connectives (e.g. "and", "however", "but") based on context. Hence, communicative

[1]Used, for example, in the BABEL generator [Goldman, 1975] to express paraphrases.
[2]Due to the modularity of GENNY, it would interesting to replace the syntactic component with a systemic grammar for comparison.

function could influence surface text, as in GENNY. However, unlike GENNY, this was done with domain dependent primitives.

Mathiesson [1980] helped develop the PENMAN generator [Mann, 1983], the largest systemic generator to date. The thrust of the research has been on the NIGEL sentence generator [Berg 1975, 1977; Halliday and Tastin, 1981 from Appelt, 1985], which converts systemic features into syntactic features using *realization procedures*.

Simmons and Chester [1982] generated sentences using bi-directional grammars in PROLOG. Rule systems which both interpret and generate language manifest a desirable property of mental models: cognitive economy. GENNY experiments with a bi-directional grammar as well as a bi-directional dictionary. However, it remains to be seen if these non-redundant mechanisms can both generate and analyze *efficiently*.

2.3.2 Lexicons

A number of researchers recognize the need for more powerful lexical mechanisms. Language entails much more than grammar and words, but linguists often avoid troubling items such as frozen phrases or conventional expressions. [Becker, 1975] suggests incorporating conventional phraseology in the lexicon since "utterances are composed by the recitation, modification, concatenation, and interdigitation of previously-known phrases." Jacobs [1985] is developing a formalism for representing a phrasal lexicon which captures both syntactic and semantic regularities in language.

2.4 When to Say it

While the work McDonald and others pursue emphasizes lexical and grammatical issues, other research has focused on the planning involved in text production. Cohen [1978] worked on planning speech acts (e.g. inform, request) in response to a user query. While his system,

OSCAR, did not generate English output, it did select an appropriate speech act, determined which agents were involved and chose the propositional content of the speech act.[1]

Appelt [1985] extended Cohen's suggestions by applying artificial intelligence planning techniques not only to speech acts but also to decisions involving syntactic structure and lexical choice. Like Cohen, Appelt viewed speech acts as communicative goals which could be modeled by planning processes. In general, he saw goal satisfaction as a complex interaction between physical and linguistic actions, ultimately motivated by the speaker's desires. Appelt implemented his ideas in KAMP (Knowledge And Modalities Planner), a hierarchical planner with multiple levels of representation including: *illocutionary acts* (request, inform), *surface speech acts* (abstract representations of the knowledge), *conceptual activation* (description selection), and *utterance acts* (surface choice).[2]

We can distinguish between the generation approaches of Appelt and McDonald as wholly pre-planned versus interleaved planning and realization, respectively. Appelt's system worked since he took into account the hearer's knowledge and state and only a limited pragmatics scope. Hence, the constrained search space made backup computationally feasible. In contrast, McDonald employed *limited-commitment planning*, allowing for two way communication between his planner and realizer. GENNY recognizes the need for flexible planning and allows informational dearth to signal to the text planner to select another message to realize. GENNY's weakness is that failure to realize a message will not signal to the planner to choose an alternate strategy. Future generators must interleave content selection, planning, and realization in a more flexible manner.

2.5 What to Say

Just as a generator must decide how and when to say something, it also must determine what to say which concerns issues of ordering, grouping, and focusing. Initial research in this

[1]Recently, Cohen [1981] proposed a planning system which determines referential descriptions.
[2]The KAMP mechanism was based on *procedural nets* [Sacerdoti 1977] which allow knowledge from many different sources to interact to solve a problem.

area [Mann and Moore, 1981] was essentially bottom-up. Mann and Moore's *partitioning paradigm* involves traversing the underlying knowledge structure depth-first to obtain grouping of propositions. While consistent for small texts, this method fails to embody the flexibility necessary to produce longer texts. The *fragment-and-compose paradigm* [Mann and Moore, 1981] does provide variability. First the message is divided into elementary propositions. Next, these are ordered using rules of aggregation (e.g. chronology). The resulting possible orderings are evaluated by means of preference values and the best organization is selected.

In contrast to these bottom-up approaches, the *text structure* view can be characterized as essentially top-down. Weiner [1980] began work on this paradigm by developing an *explanation grammar* which formalizes the ordering of propositions and characterizes text structure. Furthermore, focus of the text is controlled by a pointer to propositions throughout the explanation.

Weiner proposed that a statement can be justified by offering reasons, supporting examples, and implausible alternatives, except for the statement. These justification techniques are realized in his system by four predicates: **statement**, **reason**, **example** and **alternative**. Connectives such as **and/or** and **if/then** allow for further complexity of predicates. In order to incorporate this complexity and yet retain consistency in the surface level text, the explanation grammar rules generate trees which are further altered by transformational rules to form a hierarchical structure representing the explanation. At this stage, nodes in the tree (representing focus) are selected so as to achieve a natural flow of ideas.

These ideas were expanded and improved upon by McKeown in her TEXT system [McKeown, 1985]. Her system generates textual responses to questions about the Office of Naval Research (ONR) data base on ships. McKeown identified three types of user requests to the ONR database: requests for definitions, requests for available information, and requests for the difference between two objects. (In contrast, GENNY answers not only definitional and difference questions, but also examines generation of diagnostic explanations.)

As in previous systems, McKeown delineates the strategic (what to say) and tactical (how to say it) components of the natural language generation problem. Unlike previous systems which traced some underlying knowledge structure to generate text, her system is guided by descriptive strategies. She allows focus information provided by the message from the strategic component to influence syntactic structure.

McKeown's work is based on a formal theory of discourse strategy and focus of attention. She introduces *rhetorical techniques* which are in essence a schema or a text structure outline (figure 2.2). These aid in selecting propositions from a *relevant knowledge pool*: a source of pertinent information generated from the knowledge base. In addition, a *focusing mechanism* provides low-level coherency by connecting current and previous focuses of attention [Sidner, 1979]. Her tactical component (Bossie [1981]) translates messages into English using a functional grammar, based on Kay's [1979] formalism.

- **Requests for Definitions**
  - identification
  - constituency
- **Requests for Available Information**
  - attributive
  - constituency
- **Requests About the Difference Between two Objects**
  - compare and contrast

Figure 2.2 TEXT schema [from McKeown, 1985, p. 41].

McKeown's text generator, based on written not spoken language, has no mechanism for self-correction, ellipsis, ungrammaticality, informal phraseology, style, interruption or circularity. She has suggested that a more powerful control mechanism could augment the systems performance. In particular, a backtracking mechanism [ala Appelt, 1985] would allow for self-correction. She has recently extended her model to tailor explanations for the user [McKeown *et al.*, 1985] in an advisory system for course selection.

Just as McKeown examined generating descriptions from data bases, so Kukich [1984] developed a system, ANA, which generates stock reports from a knowledge base of daily trading on the Dow Jones stock exchange. While McKeown contributed a clear and well-motivated theory of text structure and focus, Kukich [1984] developed algorithms to obtain fluency in DB reports. She argues that attaining fluency is difficult because it relies on many different types of knowledge: semantic, lexical, syntactic, grammatical, and rhetorical. She proposed interaction between these different knowledge sources to guide surface choices. Future text generators should be guided not only by discourse strategies and fluency mechanisms, but also by models of speaker and hearer intent.

2.6 Why do we Say it?

We have seen that generators must incorporate mechanisms to determine how, when, and what to say, but more sophisticated generators must decide *why* to say it. This will require a wider range of pragmatic reasoning when producing utterances. A pioneer system, ERMA [Clippinger, 1974], attempted to model false starts, hesitations, and suppressions by incorporating a series of sophisticated modules. These included CALVIN (topic collection and filtering), MACHIAVELLI (topic organization and phraseology), CICERO (realization), FREUD (monitoring the origins of rhetorical plans), and LEIBNITZ (a "concept definition network"). While some of their functions clearly include issues addressed previously, others suggest a much broader influence on text (e.g. self monitoring).

PAULINE [Hovy, 1987] (Planning and Uttering Language in Natural Environments) can be viewed as a parameterization of ERMA. PAULINE characterizes conversational setting in terms of conversational atmosphere (the speaker, the hearer, the speaker-hearer relationship) and characterizes interpersonal goals of the hearer and the speaker-hearer relationship. For example, in a particular discourse, the speaker is represented in terms of his knowledge of the topic (expert, student, novice), interest in the topic (high, normal, low), opinions of the topic (good, neutral, bad) and emotional state (happy, angry, calm).

PAULINE represents a set of *rhetorical goals* which act as intermediaries between the pragmatics of the system (the speaker's interpersonal goals and conversational setting) and the syntactic decisions (a phrasal lexicon and syntactic experts). Thus, one can set the above pragmatic parameters to effect the rhetorical goals, ultimately realizing in stylistic English. These rhetorical goals include formality, simplicity, timidity, partiality, detail, haste, force etc. Formality, for example, can be highfalutin, normal or colloquial.

Hovy argues for this distinct level of stylistic representation since pragmatic effect is seldom the result of a single rhetorical goal but often rather a complex interaction of many (see discussion Hovy, 1987, pp. 36-38). Furthermore, rhetorical goals offer a practical (certainly partial) attempt at the problem-laden field of pragmatics. This work indicates exciting uncharted territory for further exploration.

2.7 What do we do Now?

We have summarized the origins and the current directions of natural language generation. Canned text, while as fluent as the composer, is adequate only for the most basic of applications. Furthermore, this method fails to reflect system modifications. While code conversion accounts for changes in the underlying formal representation, longer texts introduce significant coherency problems.

Recent research efforts have resulted in linguistically-motivated models from which we can build generation systems. Deciding how to say something requires a mapping of a rhetorical pattern onto surface text and can include phrasal choice, lexical selection, as well as user models. Planning when to say it should be guided by the interaction between speaker and hearer (it entails such notions as speech acts and communicative goals) to help mold text to be sensitive to the audience. Determining what to say may be interleaved with deciding when to say it and will require the use or generation of rhetorical patterns which reflect the discourse role or function of the text and the type of audience. This includes the issue of content, which should be guided by the goal and topic of the discourse (taking into account relevancy, scope and cogency). Finally,

intelligent text generation systems of the future must incorporate mechanisms for selecting words, referents, and syntax based on a user model.

We now need to investigate the pragmatic effects on surface form and the employment of devices (lexical, structural and semantic) to enhance textual connectivity and plausibility. Combining the ideas detailed in the next chapter, GENNY examines the use of some pragmatic information (e.g. focus, given/new) to constrain surface form.

Chapter 3

FUNCTIONAL LINGUISTIC FRAMEWORK

What is needed and what has been lacking, is a cohesive theory of how humans understand natural language without regard to particular subparts of that problem, but with regard to that problem as a whole.
Roger Schank

3.1 Introduction

In [Dik, 1978], the formal and the functional linguistic paradigms are contrasted. The formal paradigm (the basic view underlying Chomskian linguistics) defines a language as a set of sentences whose primary function is the expression of thoughts. In contrast, the functional paradigm defines language as an instrument of social interaction, with a primary purpose of communication. While the formal paradigm describes sentences independently of the setting (context and situation) in which they are used, Dik's functional paradigm allows linguistic expressions to be molded by their function within a given setting. Furthermore, while the formal perspective regards language universals as innate properties of humans, the functional view explains language universals in terms of the constraints of the goals of communication, the biology and psychology of the communicators, as well as the setting of the communication. In general, the relation between pragmatics, semantics and syntax within the formal paradigm is one of subservience. Within the functional framework, conversely, pragmatics influences semantics and semantics effects syntax.

The design of GENNY was guided by the functional paradigm. Provided a discourse goal, GENNY employs knowledge, discourse, pragmatic, semantic, relational and syntactic constraints to generate natural language. While the current implemented generation system by no means incorporates the whole of Dik's functional perspective (e.g. there is no analysis of the setting of the discourse, nor of the participants[1]), it nevertheless establishes a framework from within which these aspects can be investigated.

3.2 Recent Insights

This representation incorporates several recent advances in computational linguistics (and is suggestive of future extensions). These include (discussed later in detail) Barbara Grosz's [1977] ideas on global focus (knowledge relevancy, implicitly focused entities, and focus shifting), Candace Sidner's [1983] use of local focus in anaphora resolution, David McDonald's [1981] work on knowledge and message formalisms, Douglas Appelt's [1985] ideas of planning utterances, and Kathy McKeown's [1985] rhetorical-predicate based discourse model. Generative semantics in GENNY are represented in the case formalism [Fillmore, 1968, 1977] while syntax follows the GPSG [Gazdar, 1982] approach.[2] The theory of Relational Grammar, [Perlmutter, 1980], which recognizes a distinct level of grammatical primitives (e.g. subject, object) is used by GENNY to bridge a recognized semantico-syntax gap.[3]

3.3 Theoretical Overview

Figure 3.1 illustrates the levels of representation incorporated into the GENNY text generation system. The analysis begins with the knowledge representation, followed by a

[1]Although preliminary studies in this challenging area of user modelling suggest that GENNY could naturally incorporate a naive/expert distinction when selecting relevant knowledge as well as when choosing grammatical or lexical expressions.

[2]While the claim of working within a functional paradigm may seem inconsistent with the use of a Chomskian-based syntax representation, it should be noted that the unification GPSG feature grammar represents the functional analysis of a language at the syntactic level. Of course, this strata is constrained by knowledge inherited from higher levels (e.g. focus and discourse information).

[3]The difficulty of bridging the semantico-syntax gap is manifested by "hard-wired" tactical generation components.

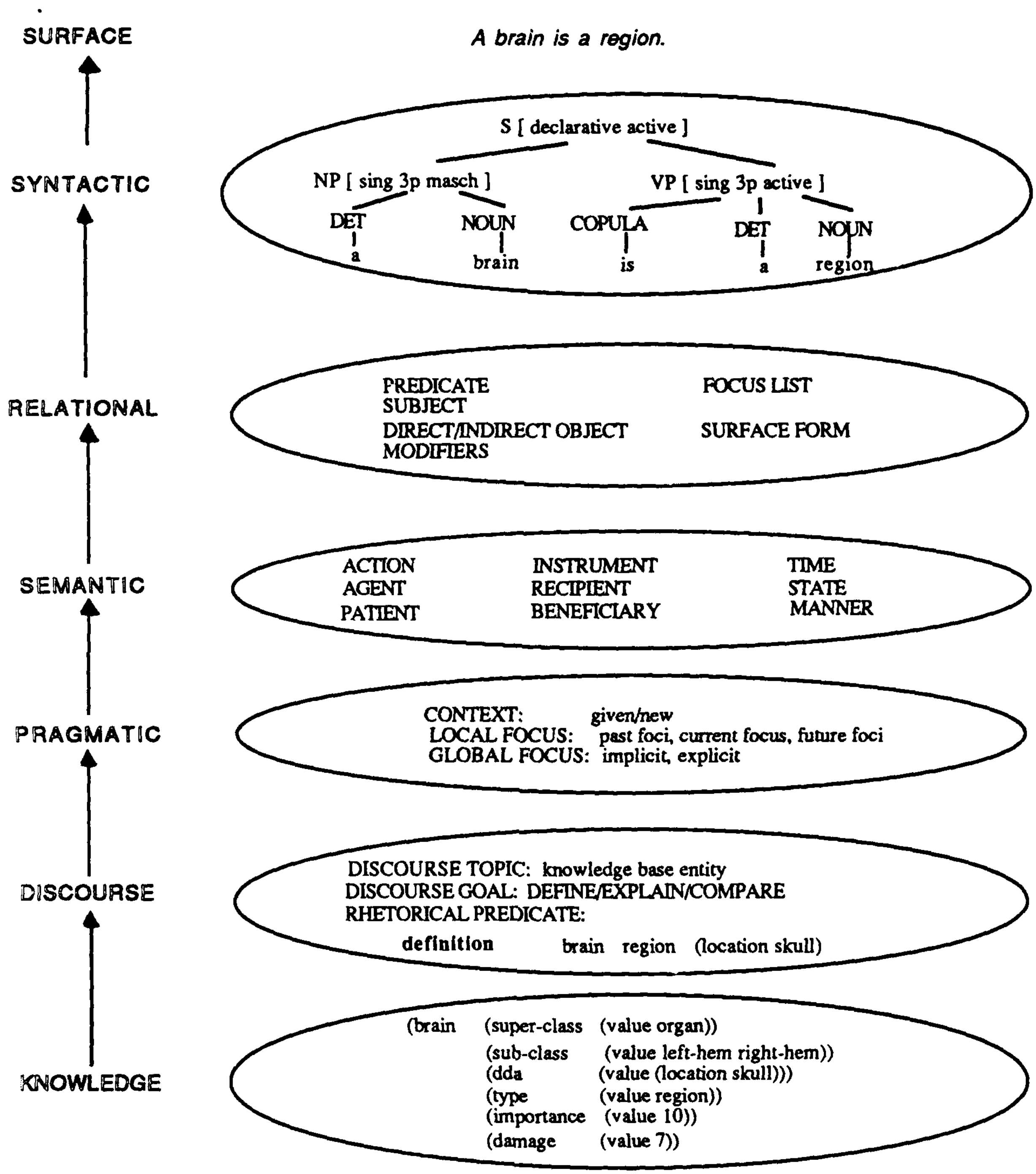

Figure 3.1 Functional Linguistic Framework

functional representation on successive levels: discourse, pragmatic, semantic, relational. syntactic and surface. The knowledge representation is the underlying method of organization used in the domain application (e.g. frames, rules). The discourse model embodies text schemas consisting of rhetorical primitives -- basic building blocks of larger texts. Interleaved with the discourse model is a pragmatic representation incorporating a focus model together with a context mechanism. A case-role semantic analysis is mapped onto a language-independent relational representation, which is then used to build a syntactic tree. Morphological and orthographic procedures generate the final surface form.

3.4 System Overview

The generator mirrors this linguistic approach. GENNY begins a session by printing (where X and Y are KB entities):

> GENNY can answer questions of the form:
>
> What do you know about X?
> Can you explain Y?
> What is the difference between X and Y?

Next, GENNY inputs a domain dictionary and knowledge base, and then queries for a discourse topic and discourse goal. Consider the session to output the first text presented in Chapter 1 (user reply in capitals):

> Please enter the domain dictionary file name?
> **NEUROPSYCHOLOGY.DICT**
>
> What is the domain of discourse?
> **NEUROPSYCHOLOGY.KB**
>
> What do you wish to speak about?
> **BRAIN**
>
> Do you wish to DEFINE, EXPLAIN or COMPARE?
> **DEFINE**

For reasons of simplicity, the discourse topic provided by the user is assumed to be the explicit name of a frame within the knowledge base. Practical generators must perform the non-trivial task of mapping user query onto knowledge base entities. In GENNY, a more plausible approach would be to perform semantic analysis on the given lexical item [c.f. Sparck Jones and Tait, 1984], which then could indicate a discourse topic(s). This is a non-trivial issue as frame selection will be problematic for a KB whose underlying representation does not parallel natural language. GENNY exploits this simplification in order to concentrate efforts on other compelling issues such as discourse structure and focus shift.

GENNY uses the discourse **topic** to generate a pool of related information (*knowledge vista)* for possible use during discourse formulation. GENNY uses the discourse **goal** to select a discourse plan (*theme-scheme*) which will guide the overall structure of the text and provide top-level cohesion. Stepping through the plan, GENNY uses global focus constraints on relevant knowledge together with local focus constraints on available propositions to select the next message (*rhetorical predicate*) to utter.[1]

Once selected, GENNY attempts to produce a message by sending it first to a semantic interpreter which maps entities onto semantic roles based on their position in the message formalism. The interpreter also exploits semantic markers which identify modifiers (e.g. location, function) which eventually become prepositional phrases. The rhetorical predicate type suggests the action, the choice of which may be constrained by the types of knowledge present in the message formalism (e.g. objects, acts, or states).

Next, the relational module uses syntactic experts -- constituent builders which utilize pragmatic (e.g. given/new), semantic (e.g. case-lexical relations) and syntactic knowledge (e.g. phrasal components) -- to produce grammatical parts such as subject, direct-object, and predicate.

[1]On the whole, the generation process in GENNY is modular and serial, except for interleaved discourse and pragmatic processing. A successful computational model should account for the behaviour of humans. The practical advantage of a serial process is the computational simplicity and comprehensibility. One major disadvantage, however, is its psychological plausibility as a mental model. Psycholinguistic studies and neurophysiological evidence indicate that the cerebral cortex exploits its highly parallel structure to solve problems concurrently [Golden, 1985].

Focus information suggests voice (active, passive) which is manifest in the ordering of relational constituents. It is at this stage that knowledge tokens in the message formalism are translated to lexical entries using the dictionary system. The rhetorical role the message plays in the overall discourse (e.g. cause-effect, illustration) may suggest particular sentential connectives ("because", "for example","therefore", etc.) which enhance low-level connectivity. Finally, a syntax tree is generated using a feature-enhanced phrase structure grammar and surface form is provided by morphological and orthographic routines.

A failed utterance (at the semantic, relational or syntactic level) will result in no output. Insufficient knowledge results in an attempt to fulfill discourse goals by alternate predicates or other possible foci. A psychologically plausible enhancement would be to maintain a minimum amount of information necessary to reply to the user's request. Ignorance should lead to an apology. GENNY's design would facilitate incorporation of such minimum informational constraints and remains an interesting area for further research.

While this appears to be a plausible model of the generation process, its ultimate success will depend on rigorous testing of all these components for multiple domains, knowledge formalisms, languages, and text types. (See chapter 10 for testing details.) The remainder of this dissertation illustrates and discusses each linguistic strata in turn including: knowledge, discourse, pragmatics, semantics, relations, and syntax.

Chapter 4

KNOWLEDGE AND DOMAIN

Get wisdom, get understanding
Proverbs 4,5

4.1 Domain of Discourse

The pragmatic motivation of GENNY rests on a desire for natural communication with a fault-diagnosis expert system in the domain of neuropsychology (NEUROPSYCHOLOGIST) [Maybury and Weiss, 1987]. Neuropsychological diagnosis is an approach to determining whether or not a patient suffers from neurological dysfunction. A typical evaluation with a patient consists of responding to verbal questions or performing perceptual or memory tasks which illuminate the behavioral condition of the patient. After collecting the empirical data (standardized test scores) and subjective data (clinical and qualitative observations), the neuropsychologist attempts to match the symptoms with particular categories of cerebral disorders.

In simulation of this process, a typical session with NEUROPSYCHOLOGIST would begin with some standard questions such as the age, health, family history (e.g. hereditary diseases) etc. This would be followed by more specific questions (user reply in caps):

```
How quickly did the condition appear?
Please type the word INSTANT, DAYS, MONTHS-YEARS, or UNSURE:
   INSTANT

Did the patient recently suffer cranial trauma in an accident?
   NO

Does the patient suffer from right hemisphere paralysis?
   WHY?
   Knowledge about L-HEMI-PARALYSIS helps determine the condition
   of the LEFT-HEMISPHERE.  Values from the RIGHT-FINGER-TAPPING
   test and the SAGGING-FACE-LIMPING-WALK observation determine the
   value of L-HEMI-PARALYSIS.  LANGUAGE, COMPREHENSION, NEGATIVE-MOOD,
   MOVEMENT-IMPAIRMENT, L-COG-FLEXIBILITY, MENTAL-CONTROL, and WRITING
   also are used to determine the condition of the LEFT-HEMISPHERE.


Does the patient suffer from right hemisphere paralysis?
    NO

What score did the patient receive on the famous faces naming test?
```

These information gathering questions have the fluency and coherency of the author, but require hand-encoding of an appropriate question for each new knowledge entry. Moreover, KB expansion or alteration will not be reflected in the user interface unless it too is updated. Also, the template listing of results often results in stilted (possibly misleading) output as in the above response to the query **WHY** (i.e. *Why does the patient suffer from r-hemi-paralysis?*).

After each response, the user is asked how sure he or she is of the test results or observation. This encourages a subjective analysis of all empirical evidence. Results of tests and observations are combined using Bayesian [Bayes, 1763] heuristics based on the weight and value of each piece of evidence. The user may elect not to answer a question by simply replying UNSURE.[1]

[1] If enhanced with more complete knowledge and descriptions, the system could be used as an interactive tutor.

Questions completed, the system issues its diagnosis:

DIAGNOSIS

The patient has a DISORDER with probability 0.8

DISORDER-TYPE	PROBABILITY
GLOBAL	0.3
FOCAL	0.8
AMNESIC	0.0

After the diagnosis report, the user can query for further explanation:

WHY FOCAL?

The patient has a FOCAL disorder with probability 0.8 because:

DISORDER	PROBABILITY
FRONTAL	0.3
HEAD-TRAUMA	0.8
STROKE	0.0
TUMOR	0.1
DEMYELINATION	0.0

WHY HEAD-TRAUMA?

The patient has a HEAD-TRAUMA with probability 0.8 because:

EVIDENCE	PROBABILITY
INSTANT-ONSET	1.0
MINOR-LTM-DAMAGE	0.9
ACCIDENT	0.5

WHY INSTANT-ONSET?

The patient has INSTANT-ONSET with probability 0.8 because you told me so.

While the explanation facilities provided in NEUROPSYCHOLOGIST[1] sufficed for the domain experts (a neuropsychologist), other users wanted to inquire about the structure of the

[1]These included the use of the keywords **why** and **how** followed by an entity in the knowledge base -- a functional notation representing the interrogative *Why does the patient have Alzheimer's disease?* in its elliptical form *why alzheimers*. Ellipsis in the functional notation occurs in the subject and the type of entity in the object. There was

underlying knowledge base. Requests of the form *Tell me about Alzheimer's disease* or *Describe the brain* were initial queries of naive users of the system (verbalized to the system designers). Furthermore, the general consensus was that explanatory diagnostic lists were functional, but unnatural.

The requirement for a describe or define facility to answer questions of the form *What is an X?* was consistent with Malhotra's [1975] finding that naive data base users often query the general contents of a data base rather than just specific values of entities contained in that data base. This mirrors the linguistic inadequacy of listing long chains of inference to explain reasoning in complex programs (e.g. planning programs [Schank and Abelson, 1977]). Listing inference chains encourages ambiguity by relying on the user to impose conceptual relationships between the listed objects.

4.2 Explanation

Explanation includes the major enterprise of collecting and presenting linguistically sufficient statistics and information. To begin, this task involves the questions: *How does the underlying KR affect the type and extent of additional information to be collected?* and *What mechanisms are necessary to collect and represent this information during system runs?*

GENNY, for example, instantiates the frame based model during the run of the expert system. Damage, test, or observation values are stored in slots associated with the appropriate frame in the brain and disorder knowledge base. In other systems -- rule based expert systems, for example -- appropriate knowledge gathering mechanisms would have to be developed. The extent of this task lies largely on the scope of the explanation. Schank *et al.* [1984ab, 1985] suggest that explanation can occur on a continuum:

making sense $\Longrightarrow$ *cognitive understanding* $\Longrightarrow$ *complete empathy*

---

also direct inquiry of a specific entity of the brain model (e.g. **how-bad left-frontal**) as well as a **why-useful** function for explanation, tutoring, or system debugging.

Current artificial intelligence technology deals with the lower end type of explicit explanation. A more interesting task (far beyond the scope of this dissertation) is the explanation of anomalous situations which are key to learning. [Kass and Leake, 1987] offer a categorization of explanation for intentional actions, material anomalies and social anomalies. Explanation raises issues on the frontiers of knowledge and language and, ultimately, may prove to be the most interesting (and difficult!) task or generators of the future.

4.3 Distinguishing Descriptive Attributes

Unfortunately, current knowledge representations (KR) are generally ill-suited for even the simplest of linguistic tasks, much less sophisticated explanations. Because of the hierarchical structure of the domain, frames [Minsky, 1975] were the natural method of encoding expert knowledge in the original knowledge based system. Figure 4.1 illustrates a typical frame characterizing neurophysiology as it appeared in the original knowledge base after diagnosis.

A frame-based representation is convenient, efficient and powerful. Frames consist of frame names, slots, facets, and values. In figure 4.1 the parentheses separate the different categories. The frame name is BRAIN. The different slots are SUPER-CLASS, SUB-CLASS, TYPE, IMPORTANCE, and DAMAGE. The facets of the slots in this example are all VALUE. An alternate facet name, for example, is DEFAULT. The actual values are the symbols which appear after the word VALUE in each line. The frame hierarchy is defined by the values in the SUPER-CLASS and SUB-CLASS slots. Frames which are instantiations of a particular frame TYPE inherit properties of their general frame. Importance is the relative significance of a piece of knowledge with respect to its siblings in the knowledge hierarchy.

```
(BRAIN  (SUPER-CLASS  (VALUE HUMAN))
        (SUB-CLASS    (VALUE LEFT-HEMISPHERE RIGHT-HEMISPHERE))
        (TYPE         (VALUE ORGAN))
        (IMPORTANCE   (VALUE 10))
        (DAMAGE       (VALUE 5)))
```

Figure 4.1. Top-level NEUROPSYCHOLOGIST frame

Figure 4.2 illustrates the same frame as it appears in GENNY.[1] Note the extra slot name **DDA** for distinguishing descriptive attribute [McKeown, 1985]. This is the only addition to the KB for generative purposes. The DDA (attribute-value pairs) is an additional slot in the frame which describes the justification for a hierarchy partition at this level (related to [Lee and Gerritsen, 1978] partition-attributes). In figure 4.2 the brain frame can be linguistically distinguished from other parts of the body (e.g. heart, lungs) by noting that its primary function is understanding and that it is located in the human skull. The DDAs in GENNY are more flexible than those in TEXT [McKeown, 1985] as they permit lists of values to be assigned to a particular attribute, much as frames allow lists of values for a particular slot name.

```
(brain  (super-class   (value human))
        (sub-class     (value left-hemisphere right-hemisphere))
        (type          (value organ))
        (dda           (value (location skull human) (function understanding)))
                       (importance  (value 10))
        (damage        (value 5)))
```

Figure 4.2. Illustration of GENNY frame

The knowledge base was augmented by three DDA attribute types: **function, location** and **instrument**. Of course these are only three alternatives from a large number of semantic markers which could be used to discriminate entities [Sparck Jones and Boguraev, 1987]. In fact, subsequent experimentation with a second knowledge base (photography), indicated a need for a fourth attribute, **external-location**, in contrast to a membership or internal location. These attributes were used by the semantic interpreter to assign proper roles to their values in the deep case structure. According to their analysis, these attributes eventually translate to surface modifiers. Thus, external-location might realize as "on" whereas location as "in", function as "for", and instrument as "with". The value of the attribute eventually translates to a noun phrase.

[1] Do to the large size of the KB, only representative frames (37 from 142) were actually used for generation purposes. They were carefully chosen to reflect the full range of knowledge and relationships in the original expert system.

4.4 Discussion of Linguistic and Extra-linguistic Knowledge

The fact that the DDA is the only addition to the KB suggests the suitability of a frame representation for generation purposes. This claim is supported by experiments with a second, photographic KB and lexicon demonstrating domain independence. Consider a typical output, in response to the simulated query *What is photography?*:

```
Photography is an art-form for recording images on film.
It has a relative importance value of ten.
It contains three faults:  an equipment fault, a technique
    fault and a style fault.
The equipment fault has a relative importance value of three.
The technique fault has a relative importance value of four.
The style fault has a relative importance value of nine.
It, for example, is a fault with personal expression.
```

It is fair to compare GENNY to the TEXT system [McKeown, 1985], which produced similar quality text for equivalent definitional discourse goals (although GENNY also investigated explanations). (See Chapter 10 for a comparison.) However, in addition to DDAs, McKeown found the need to augment her underlying KR (an entity-relationship DB model [Chen, 1976]) with both a *generalization hierarchy* and a *topic hierarchy*.

Under closer scrutiny, we find that the generalization hierarchy describes relations of entities (e.g. part-whole) while the topic hierarchy describes relations of attributes (e.g. type-instantiation). These additional knowledge structures would, unfortunately, have to be hand-encoded for each new formalism. While this application dependence is undesirable, it appears that there is a certain amount of additional linguistic or real-world knowledge (e.g. DDAs) which unavoidably will have to be tailor-made for each KR.

The frame paradigm, however, minimizes customization. This becomes clear when we notice that two types of relationships are being encoded in these formalisms: part-whole and type-specialization (also referred to as a-kind-of). In the frame KB, the slots named **super-class** and **sub-class** represent the part-whole relationship (classes and elements, parts and components or events and sub-events). The slot named **type** represents the type-specialization relationship

(object/entity-types and instantiations). For example, the frame in figure 4.2 encodes that a brain is a part of the human body via the **super-class** slot and that the brain is a particular type of organ via the **type** slot. This example demonstrates the clarity of the frame KR.

This raises the question as to what is the most effective KR from both knowledge and linguistic perspectives? As outlined in section 2.3, McDonald [1981] investigated a variety of KR in his text generation system, MUMBLE, including predicate calculus, PLANNER-style [Winograd, 1972] data base assertions, OWL, FRL and KL-ONE. His research suggests that different linguistic phenomena are more naturally represented in some message formalisms rather than others. OWL [Hawkinson 1975 in McDonald, 1981], for example, specifically allows codification of NL phenomena (ambiguity, quantification, etc). However, one still has the problem of interfacing this to the underlying application KR. His contribution is a **message formalism** independent of the underlying representation.

Unhappily, current knowledge formalisms are inadequate. Frames [Minsky, 1975] as well as scripts [Schank and Abelson, 1977] are difficult to select in a well-motivated manner. (GENNY avoids this problem by having the user select a frame or frames.) Furthermore, they deal poorly with non-standard objects or events. Scenarios [Sanford and Garrod,1981], which describe the "extended domain of reference", suffer similar problems with control of inference. Johnson-Laird [1983] describes knowledge in terms of a model-theoretic semantics of possible states of affairs in time and in space: mental models. The practical details of such a representation, however, remain elusive, and make assessment virtually impossible. Nevertheless, it appears suspect to the same problems as above. All KR have difficulty selecting relevant knowledge -- a problem partially addressed by global focus algorithms in GENNY (see section 6.2) but which requires further investigation.

One solution to these formalism deficiencies is to maintain two levels of representation for discourse: a superficial propositional format similar to linguistic form coupled with a mental model representing the structure of events or knowledge in the real world [Johnson-Laird, 1983]. GENNY can be viewed is this light since the frame KB models the domain as it exists structurally

and functionally in nature (a sort of "static mental model") while the rhetorical predicate level, is more closely aligned with linguistic form.[1] A predicate semantics connects the mental level to the propositional level, which serves as the basis for discourse representation to which we now turn.

[1]Psychologists believe that semantic (long term) memory in humans plays a dual role: representing the current state of our past experience of the world and forming the basis of linguistic acts [Greene, 1975: 132]

Chapter 5

DISCOURSE THEORY

If a question can be put at all, then it can also be answered.
Ludwig Wittgenstein

5.1 Introduction

Given that humans have some mechanism for storing knowledge (say in a frame-like representation), how is it that we are able to communicate effectively in response to a request for information? Humans appear to exploit standard strategies to organize and present ideas. In this chapter, we first examine what properties make a string of sentences a coherent, plausible and connected *text*. Next we examine the issues of text structure including *story grammars*, *text grammars* and *text schema*. GENNY's higher level text formalism is then presented: *theme-schemes*. This chapter concludes by discussing *rhetorical predicates*, the basic primitives of text structure.

5.2 Text

We first distinguish between written and spoken discourse as representing a divergence in functional emphasis: the former is predominantly transactional while the latter is mainly interactional. To constrain our task, we choose to focus on written text since spoken discourse

contains many interesting but difficult phenomena such as phonological idiosyncrasies and speech errors (e.g. slips of the tongue).[1] The Concise Oxford English Dictionary defines 'text' as:

(1) original words of author as opposed to a paraphrase or commentary on them.
(2) a passage of scripture quoted as authority especially as chosen as subject of sermon etc; subject, theme.

More suggestive is the definition of "texture":

> arrangement of threads etc. in textile fabric, characteristic feel to this; arrangement of small constituent parts, perceived structure; representation of structure and detail of objects in art; quality of sound formed by combining parts.

Perhaps this characterization led Halliday and Hasan [1976, p. 2] to state that "a text has texture and this is what distinguishes it from something that is not a text ... the texture is provided by the cohesive relation." This connective relationship manifests itself in text when interpretation of an utterance presupposes knowledge of a previous utterance. For example, a cohesive relation can exist as an anaphor:

Never hold onto the punt pole if it gets stuck in the mud

where the pronoun "it" refers to the preceding definite noun phrase "the punt pole." In addition, discourse can be connected with *cataphora* (forward reference) and *exophora* (extra-textual reference). Utterances can also be unified through formal markers such as "and", "however", "for example", and "then":

If you fall in the river then you will catch cam fever.

Several grammarians have classified connectives [Quirk and Greenbaum, 1973; Halliday and Hasan, 1976]. Halliday [1985, p. 302-307] offers a taxonomy of such markers: elaboration, extension, enhancement. Extension, for example, can be additive (and, also, moreover, in addition), adversative (but, yet, on the other hand,however) or variation (on the contrary, apart from that, alternately). He relates surface forms with these connectives, illustrating their cohesive function in discourse.

---

[1] We thus explicitly exclude such effects as phonology, intonation, dialect, and accent, and implicitly avoid phenomena such as spatial context (e.g. body gestures).

Clearly connective relation of text can be implied rather than explicit, such as in poetry [Johnson-Laird, 1983, p. 377]:

Swiftly the years, beyond recall
Solemn the stillness of this spring morning.

Connection is also implied in a list of historically significant dates or --as in the original explanation procedure in NEUROPSYCHOLOGIST -- as a list of possible disorder candidates.

Johnson-Laird [1983] distinguishes between the *coherence* and *plausibility* of discourse. Analyzing the response time of humans to a set of psycholinguistic experiments involving referential continuity, Ehrlich and Johnson-Laird [1982] established coherency as a property of discourse. However, they characterize plausibility as reflecting the ability to place the actual sequence of events into a temporal, causal, or intentional framework.

Clearly many devices aid the cohesion of text including co-reference, lexical relationships (hyponymy, part-whole, collocability), structural relationships like clausal substitution (e.g. "so am I"), syntactic repetition, consistency of tense and stylistic choice [see Quirk and Greenbaum, 1973, pp. 284-308]. Halliday and Hasan [1976, p. 229] claim that the heart of cohesiveness "is the underlying semantic relation." Hobbs [c.f. Carter, 1985] provides the noteworthy distinction between *coherence*, which stems from the conceptual relevance of the text content, and *cohesion*, which arises from textual linkages.

5.3 Story Grammars

It was precisely this textual connectivity that Rumelhart and others attempted to capture in *story grammars*. These grammars codified stereotypical scenarios, found in genre such as folk tales, into content-independent structures in the same spirit that grammarians captured regularities in syntactic structures. Figure 5.1 illustrates a simple example with both syntactic and semantic rules [Rumelhart, 1975 from Johnson-Laird, 1983, p. 363]. The greatest weakness in the story grammar formalism is its lack of specificity: terminal categories lack explicit definitions and semantic rules rely heavily on world-knowledge.

The syntactic rules

1 Story	->	Setting + Episode
2 Setting	->	(State)* [i.e., an arbitrary number of states]
3 Episode	->	Event + Reaction
4 Event	->	{Episode Change-of-state Action Event+Event}
5 Reaction	->	Internal response + Overt response
6 Internal response	->	{Emotion Desire}

The semantic rules (corresponding to each syntactic rule)

1 Setting ALLOWS episode, i.e., makes it possible.
2 State AND State AND ..., i.e., logical conjunction of the states.
3 Event INITIATES reaction, i.e., an external event causes a mental reaction.
4 Event CAUSES event, or event ALLOWS event. (No semantic rule is require for the first three options in the syntactic rule.)
5 Internal response MOTIVATES overt response, i.e., the response is a result of the internal response.
6 No semantic rule required.

Figure 5.1 Rumelhart's Story Grammar from Johnson-Laird, p. 363.

These formalisms had some utility, namely the classification of repetitive stories. For example, they could capture the repetitive style of the biblical story of genesis which essentially follows the pattern:

DayN -> Divine-suggestion + object-creation-event + object-naming +"Evening came and morning followed, the nth day."

Figure 5.2 shows an abbreviated form and translation of a popular Italian folk-song which can be interpreted by the story grammar because of its regular recursivity. As this example illustrates, the power of a context free grammar is unmotivated since a finite state machine which allowed for say 100 repetitions of the *event -> event + reaction* rule would suffice for all stories with this structure. In sum, these indefinite rules were a contribution, but lacked descriptive precision. More importantly, they were text-type dependent.

Alla Fiera Dell'Est	At the Eastern Fair
Alla fiera dell'est	At the Eastern fair
per due soldi	for 2 pieces of money
un topolino	a little mouse
mio padre comprò	my father bought
E venne il gatto	And then came the cat
che si mangiò il topo	that ate the mouse
che al mercato	that at the market
mio padre comprò	my father bought
E venne il cane	And then came the dog
che morse il gatto	that bit the cat
che si mangiò il topo	that ate the mouse
che al mercato	that at the market
mio padre comprò	my father bought
...	...
...	...
E in fine il Signore	And in the end God
sull'angelo della morte	on the angel of death
sul macellaio	on the butcher
che uccise il toro	that killed the bull
che bevve l'acqua	that drank the water
che spense il fuoco	that extinguished the fire
che bruciò il bastone	that burnt the stick
che picchiò il cane	that beat up the dog
che morse il gatto	that bit the cat
che si mangiò il topo	that ate the mouse
che al mercato	that at the market
mio padre comprò	my father bought

Figure 5.2. Angelo Branduardi's Alla Fiera dell'est

5.4 Text Grammars

The solution to non-specificity of grammatical rules was a domain dependent representation of discourse: *text grammars*. This is illustrated by the top level text grammar rule in a generation system for a neurological data base for strokes [Li *et al.*, 1986]:

Case-Report -> Init-Info + Md-Hsty + Fin-Dex + Phy-Exam + Lab-Tst + Outcome

Expanding the fourth constituent of this rule we get:

Phy-Exam -> General-Exam + ... + Cerebellar-Exam

Note that the terminal and non-terminal categories are domain dependent. Also, the Li *et al.* generation system is KR dependent. In contrast, GENNY maintains a linguistically independent representation of the underlying knowledge: rhetorical predicates. Predicate semantics link these linguistic primitives to GENNY's KR. The rhetorical primitives formulate the basis for text schema, a discourse formalism independent of domain and text type.

5.6 Text Schema

Making reference to Plato's visualization of the true triangle, Kant [1787] writes "In truth, it is not images of objects, but schemata which lie at the foundation of our pure sensuous conceptions." Recent studies by the cross-cultural psychologist Elanor Rosch [1976] demonstrate psychological evidence that natural categories are represented in prototypes. These and other arguments lend credence to the philosophical and psychological adequacy of representing discourse in *schema*. Their empirical success is the nature of this dissertation.

Perhaps the first instigator of text schemas was Aristotle who distinguished between two discourse techniques: enthymemes (syllogisms) and examples. Enthymemes are types of arguments; examples support these arguments. But just as story and text grammars suffer from generality, so these broad categories offer little insight into the cohesive relation of utterances within a multi-sentential text.

Of late, grammarians [Williams, 1893; Scott, 1938] have categorized the function of paragraphs in text as "topic, general illustration, particular illustration, comparison, amplification, contrasting sentences, and conclusions." Grimes detailed this to describe rhetorical predicates as serving an organizational function in discourse [Grimes, 1975]. Accordingly, predicates can support or supplement, locate (spatially or temporally), and identify. Searle [1969, 1975] noted that using the wrong rhetorical predicate to purposefully flaunt the maxim of relevancy will cause a conversational implicature.

McKeown examined the ordering of these communicative techniques by analyzing text produced by humans. She developed several schema which represented sequencing of predicates

{Identification (description of an object in terms of its superordinate)}[1]

Attributive* (associating properties with an entity) /Cause-effect*

Constituency (description of subparts or subtypes)

{Depth-identification / Depth-attributive

{Particular Illustration / Evidence}

{Comparison ; Analogy} }+

{Attributive / Explanation / Analogy}

Figure 5.3 Constituency schema in TEXT [from McKeown, 1985, p. 41].

to achieve a particular discourse goal: **attributive, identification, constituency, compare** and **contrast**. She organized these into descriptive and comparative strategies (figure 5.3).

McKeown's work followed work on text grammar by van Dijk [1977], who argued that mere co-reference in text was not sufficient for producing well-formed discourse. Van Dijk suggested "macro rules" which, guided by a scheme representing the speaker's goals, could express propositions based upon their relevancy to the discourse topic. Moreover, his work suggested an interpenetration of linguistic and factual knowledge which implies that both a sense (propositional model) together with a significance (mental model) formalism are at work. Such models could well prove to be the cohesive framework of text (such as 'plot' in narrative, 'topic' in non-fiction, etc.).

[1]Using McKeown's notation: "{} indicates optionality, "/" indicates alternatives, "+" indicates that the item may appear 1 or more times, "*" indicates that the item may appear 0 or more times, and ";" suggests that the propositions can not be clearly classified as corresponding to one predicate.

5.7 Theme-Schemes

GENNY embodies an attempt to explicitly formulate and utilize common discourse strategies found in human produced text. While the discourse approach is similar to the work of McKeown [1985], the formulation is motivated by unique discourse requirements, namely that of providing definitions and comparisons of the knowledge and explanations of the reasoning within an expert system for neuropsychological diagnosis. Moreover, there exists a clearer distinction in GENNY than in TEXT of the "mental model" and the "propositional format". For example, the *message formalism* in GENNY, *rhetorical predicates* (discussed in the next section), is more linguistically independent. No linguistic markers such as restrictive or non-restrictive clauses appear in the message formalism, only semantic indicators (DDAs). Propositional content derives directly from the frame knowledge base used in the NEUROPSYCHOLOGIST expert system. In contrast, McKeown hand[illegible]d a knowledge base that represented the ONR data base about sea-going vessels. Finally, the translation from rhetorical predicates to surface form proceeds via a series of modular transformations which access explicit knowledge of semantics, grammatical relations, and syntax.

A *theme-scheme* uses the message formalism to build text types. As in McKeown's [1985] system, a text consists of standard sequence of rhetorical predicates found to occur in natural text. Rhetorical predicates classify the rhetorical function that a piece of text (sentence or clause) performs within the larger linguistic framework (theme-scheme). Predicate groupings are not necessary and sufficient for well formed text, but typical. Text from magazines, books, and advertisements were analyzed in search of common organizational strategies. Consider paragraph two from the forward of the Cambridge University *Varsity Handbook* [1986]:

> The Varsity Handbook is different. It does not attempt to present a unified and neatly packaged version of the 'real' Cambridge. It is written and produced entirely by students and reflects a range of opinions. The 'University' section is an assortment of articles by students on aspects of University life. The 'Time Out' section is intended to suggest ideas about how to spend your

```
spare time in and around Cambridge and includes an extensive
restaurant and pub guide.  The 'Information' section is a useful
file of the many services and facilities available in the area.
```

Note how the text first defines the handbook, tells about some of its attributes (what it is and what it is not), and then introduces each of its constituent parts in turn.[1] From similar analysis on many examples, the following frameworks of ordered rhetorical predicates were abstracted:

DEFINE	**EXPLAIN**	**COMPARE X,Y**
definition	cause-effect	definition X
attributive	attributive*	attribute X
constituent		definition Y
attributive*		attribute Y
		compare-contrast X Y
		inference

But with subsequent examination, a separate level of abstraction was discovered: **sub-schema**. These can be viewed as the sub-acts which are employed to realize a rhetorical act such as define, explain or compare:

THEME-SCHEMES

DEFINE	**EXPLAIN**	**COMPARE X,Y**
introduction	reason	introduction X
description	evidence	introduction Y
example		comparison X,Y
		conclusion

[1] As Schank (1977) points out, people consistently leave out redundant or obvious information to be more concise. Anaphora, for example, indirectly refer to something at the forefront of the discourse. Omission of connectors in causal chains are a similar phenomena. In the extract, notice the suppression of the sentence introducing the sections in the *Varsity Handbook*.

SUB-SCHEMA

introduction	->	definition + attributive
example	->	illustration
description	->	constituent
constituent	->	attributive* | definition*
conclusion	->	inference
reason	->	cause-effect
comparison	->	compare-contrast
evidence	->	attributive* | definition*

For example, in response to the simulated request *Why do you think the patient is unstable?*, GENNY would explain:

```
The instability symptom is manifest because the personality
observation and the sex-activity observation indicate damage. The
personality observation has a likelihood value of four. The sex-
activity observation has a likelihood value of four.
```

Text analysis uncovered other informational constructs such as *persuasion -> position / statement + justification*. Also, cause-effect predicates are often reversed. (See the Appendix for a detailed example.) For effective text, however, we must realize these higher level acts under pragmatic constraints.

5.8 Rhetorical Predicates

The basic building blocks of discourse, *rhetorical predicates* (RP), describe the relative communicative role an utterance plays within a discourse. The nomenclature for RP in GENNY arises from their function within the thread of discourse including: definition, attributive, constituent, evidence, illustration, cause-effect, compare-contrast, and inference. The selection of a particular RP is motivated by the theme-scheme employed.

A RP is instantiated with information from the KB, having been provided an argument which represents the current discourse topic entity or focus of attention (corresponding to a frame in the KB). Furthermore, this can depend upon the type of discourse goal as with the attributive RP which will be instantiated with damage information if the discourse goal is DEFINE, with

importance information if the goal is EXPLAIN, and with both if the goal is COMPARE. This is interesting because the rhetorical predicate content relies not only on its role in discourse but also on the type of discourse structure involved.

A *predicate semantics* is defined which relates the entities, relations, and values in the KB with the appropriate RP slots. For example, given the RP type, 'definition', together with the discourse topic entity, 'brain', the predicate instantiation routine returns a message with the entity, superclass, and DDA.

```
(definition ((brain))
            ((organ))
            ((location (skull human)) (function (understanding))))
```

Depending on the context (such as the past focus of information as well as the amount of given/new information) the message could eventually be realized as *A brain is an organ for understanding located in the human skull.* The predicate semantics are domain independent, as illustrated by generation from two knowledge bases (brain and photography faults). The semantics are knowledge representation specific and would have to be redefined if, for example, a script [Schank and Abelson, 1977] formalism replaced the frame KB. Given the system modularity, the amount of programming effort would be minimal. The complete predicate semantics are documented in [Maybury, 1987b, volume II, section 6].

While GENNY'S theme-schemes and their corresponding rhetorical predicates model common discourse strategies employed by humans, these alone will not generate well-connected and plausible text. Humans use knowledge of focus of attention as well as knowledge of context to decide what to utter. In this light,the selection and realization of the RP is constrained by pragmatic information, which we now discuss.

Chapter 6

PRAGMATICS

*Without knowing the **force** of words it is impossible to know men.*
Confucius, Bk XX, 3

6.1 Introduction

When reviewing the pragmatics literature one state of affairs becomes immediately evident: terminological chaos and inconsistency. To begin with, the scope of pragmatics itself is ill-defined. Oversimplifying, it includes the communicators' identities, their knowledge, intentions and beliefs, as well as the temporal and spatial setting of the speech act: context. Pragmatics has been contrasted with grammar (in the broad sense incorporating phonology, syntax, and semantics):

> *[Grammars] are theories about the structure of sentence types ... Pragmatic theories, in contrast, do nothing to explicate the structure of linguistic constructions or grammatical properties and relations ... They explicate the reasoning of speakers and hearers in working out the correlations in a context of sentence tokens with a proposition. In this respect, a pragmatic theory is part of performance.* [Katz, 1977, p. 19]

But clearly some contextual features effect grammatical structure. We select the passive over the active voice to stress what is normally the object by promoting it to the subject position. Consider:

(a) *John hit Mary with the stick.*

(b) *Mary was hit by John with the stick*

We select (b) to emphasize Mary. If we want to emphasize that John (not Mark) hit Mary we could use extraposition (*It was John who hit Mary*), or intonational stress (***John** hit Mary*).

In fact the opposite of the grammatically-based view states that pragmatics is the interaction between language and context which yields particular grammatical structures. While this perspective includes the study of deixis (extra-textual reference such as "this" or "that"), presupposition, and speech acts, it would unfortunately exclude conversational implicatures, as they are non-grammatical. Its virtue is the clear delineation and exclusion of sociolinguistics and psycholinguistics. But this pragmatic-grammatical link seems tenuous for when a Peruvian immigrant speaks English with a heavy South American accent, it is more than likely not the result of a correlation between linguistic form and context. On the contrary, this phonological eccentricity, as with a drunk's slur, is unintentional. However, selecting the Italian "tu" verb conjugation when speaking to a lover on the back of a gondola near the Piazza San Marco *is* a pragmatic-driven grammatical choice.

Since the greatest weakness of this last definition is the lack of coverage of extra meaning (e.g. implicatures), we are led to Gazdar's [1979, p. 2] formulation:

> Pragmatics has as its topic those aspects of the meaning of utterances which cannot be accounted for by straightforward reference to the truth conditions of the sentences uttered. Put crudely PRAGMATICS = MEANING - TRUTH CONDITIONS.

Because of the complexity (and excess baggage) of the term 'meaning', Levinson [1983, p. 14] sidesteps the definition and instead describes the communicative content of an utterance as including truth conditions or entailments, conventional implicatures, presuppositions, felicity conditions, conversational implicature (generalized and particularized) and inferences based on conversational structure. To be sure, pragmatics includes the Gricean cooperative principle as well as the maxims of quality, quantity, relevancy, and manner [Grice, 1975].

Clearly no current text generator comes near to providing all or even a significant subset of these capabilities (but see [Hovy, 1987]). GENNY's pragmatic analysis is confined to one Gricean maxim: be relevant. GENNY generates under the pragmatic constraints of global focus of

attention, local focus of attention, and current context (previously uttered entities). GENNY uses the topic of the discourse to globally select relevant information from the KB, local focus information to select from among alternative rhetorical predicates as well as to choose relational structure, and current context information to decide on referring expressions as well as to guide lexical choice. Working together, these constraints contribute to a more connected and cohesive text.

6.2 Global Focus -- The Knowledge Vista

Once the user has provided the topic of discourse,[1] a specific entity within the frame KB, a vista of relevant knowledge (*kvista*) is generated. This is motivated by Grosz's [1977] focus theory. Essentially, knowledge relevancy is the knowledge equivalent of phonological stress whereby entities in the KB are distinguished as being explicitly, implicitly, or not at all in focus. Figure 6.1 represents global focus in operation in GENNY.

Entities explicitly in global focus throughout a discourse are those objects tightly coupled to the discourse topic. In GENNY, this includes the discourse topic frame itself, its parent frame(s), and its child(ren) frame(s). Frames less salient, but still closely connected to the discourse topic, are placed in implicit focus. GENNY includes siblings (frames on the same level of the hierarchy) in this focus group. All other frames are not globally focused. Thus frames with a super/sub-class relationships (part-whole) are placed in explicit or implicit focus based on their distance from the discourse topic frame. It could also be argued that frames of the same type as the topic focus could be placed in focus (i.e. of the same value in the slot "type"). However, it seems implausible that since the left-hemisphere region is focused, all other frames of type "region" should be focused.

[1]The overall focus of attention for text. It should be all things related to this. The simplest, and non-trivial case, is where it actually corresponds to an actual entity within the KB.

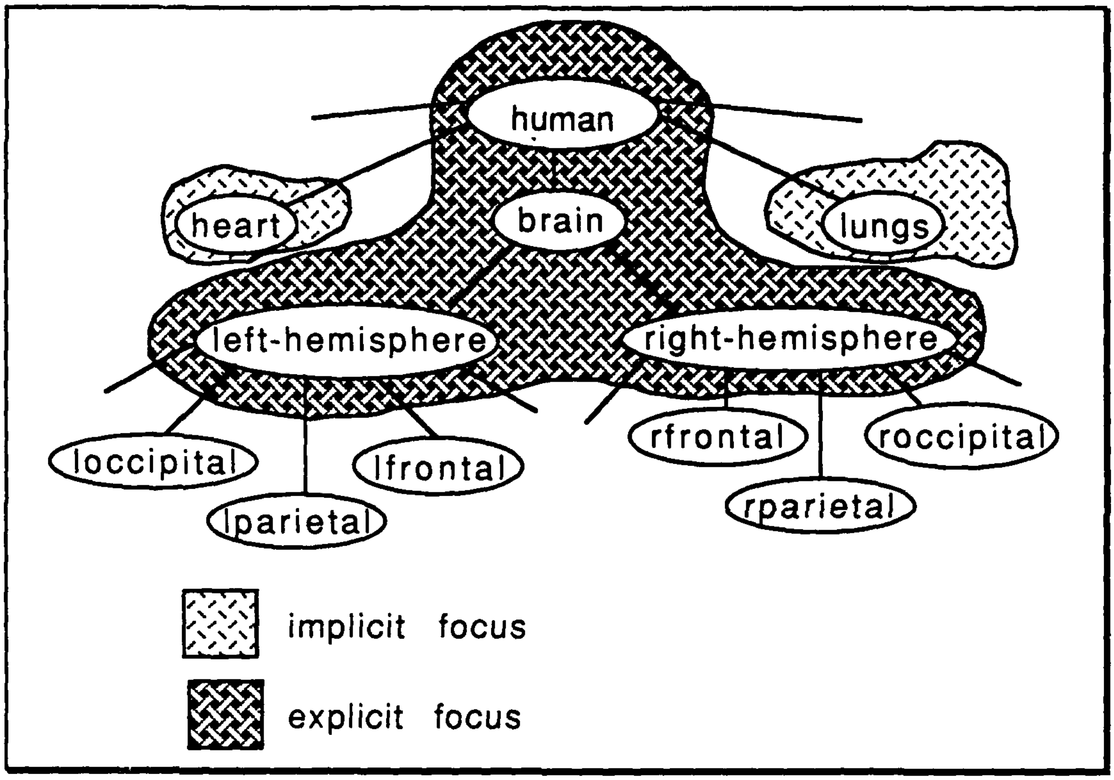

Figure 6.1 Illustration of explicit and implicit focus in KB.

Hence, from a global perspective, knowledge is viewed as entirely, partially, or not at all relevant. Of course this defines a vista with respect to the level of detail of relevant knowledge. Another powerful mechanism would be the global focusing of individual slots on information guided by the particular overall view or perspective of things. So that while the brain frame might be in explicit global focus, the kvista on the domain could determine the relevancy of functional versus structural knowledge constrained by the overriding perspective (see discussion of [Hendricks, 1975] imposing visibility constraints in [Grosz, 1977]; [McCoy, 1985]).

6.3 Focus Shift

Just as discourse tends to center on one topic, so too conversation is governed so that it flows naturally from one idea to the next. Humans change focus locally (from utterance to utterance) by either direct locution as in "We have finished our discussion about X and will now turn to Y" or by implicit means, as in "Anyways, how is the weather?" Intuitively, there are

"open" foci, in the sense that they can be mentioned without considerable worry of connectivity as well as "active" foci, which seem even more at the forefront of our minds [Grosz, 1977].

Two general principles seem to govern focus shift in discourse [Brown and Yule, 1983, p. 67]. *The Principle of Analogy* holds that things tend to be as they were before. *The Principle of Local Interpretation* claims that if there is a change, assume it is minimal. Assuming the Gricean principle of cooperation, humans exploit these discourse principles and other coherence cues when interpreting text. Unfortunately, these vague terms beg for concreteness within a computational model of generation. I define focus as:

> *something placed at the forefront of our mind by implicit or explicit means, by grammatical constructs or phonological stress.*

Three types of focus (motivated by [Sidner, 1983]), operating at the utterance level, are recognized in GENNY: current focus (CF), past focus (PF), and future focus (FF). I define:

CF -- generally the semantic actor, the subject of the sentence, the leftmost np of the sentence, and given.
PF -- past foci stack -- simulates a long-term, multi-utterance episodic memory
FF -- generally semantic patient, object of the sentence, residing at the end of the sentence new information

McKeown [1985] exploited insights made by Sidner [1979], and controls focus choice by preferring potential future foci to current focus as well as preferring current focus to the past current focus. A final alternative allows her to choose semantically related entities.

If we blindly follow the linguistic principle of analogy, our preferred choice of subsequent focus should be **CF > FF > PF** in the current discourse (were ">" means "is preferred to"). Of course with this approach speakers would drone on about one subject until exhausting their knowledge or energy.[1] In ordinary discourse, however, speakers tend to shift to recently introduced or new entities found in the future foci of the previous utterance. This suggests a promotion of FF in our rule to obtain a focus preference function: **FF > CF > PF**. If there are multiple CF (as when comparing objects), however, we should encourage discussion of those

[1]This would perhaps be a useful strategy in some situations (e.g. filibuster during a congressional session, attempt to bore at a party, or simulating a one-track mind).

before moving on to new topics. This is reflected in GENNY preferring **CF > FF > PF** when there are multiple CF. This focus preference list (fpl) plays a key role in the attentional algorithm and predicate choice in GENNY (see figure 6.2).[1] A trace of the focus selection algorithm in action is presented the Appendix.

Speakers are often encouraged to "stick to the point", which would suggest a constraint on the proliferation of new foci of attention. GENNY is discouraged from straying away from the discourse topic by knowledge vista constraints which limit the knowledge base available for discourse construction. Furthermore, when GENNY runs out of new things to say, she can always return to the original topic of discussion as it will be the first item to be placed on the PF stack.

[McKeown, 1985] suggests the need for an additional focal selection for implicitly related entities. In GENNY, a global focus of attention places related entities into the knowledge vista. Interestingly, interfacing to a rule-based representation would require a global focus routine with semantic knowledge of related entities.

Of course a more sophisticated memory device for past foci might have decay register whereby with time (say measured by the number of utterances produced), previously focused entities fade away from the forefront of discourse. In addition, a spreading activation mechanism (similar to that employed in CAPTURE [Alshawi, 1983]) could encourage frames that are related to the current focus of attention to become more strongly in focus as they are spoken about or referred to. A provocative idea would be to use the amount (and strength) of KB links to the current focus to suggest future foci. Local and global focus mechanisms remain an exciting area for further research.

[1] Fillmore [1977] suggests that entities in an event are *perspectivised* and claims a need for a saliency hierarchy -- a priority list of foreground choices which can be used to decide on focus. He suggests an animacy hierarchy can aid perspectivisation decisions. Given a choice, egocentric people tend to focus first on humans, then animate things, and finally on inanimate objects. Animacy knowledge could easily be added to lexical entries and GENNY's focus algorithm could be adapted to make such decisions. There was no time for implementation.

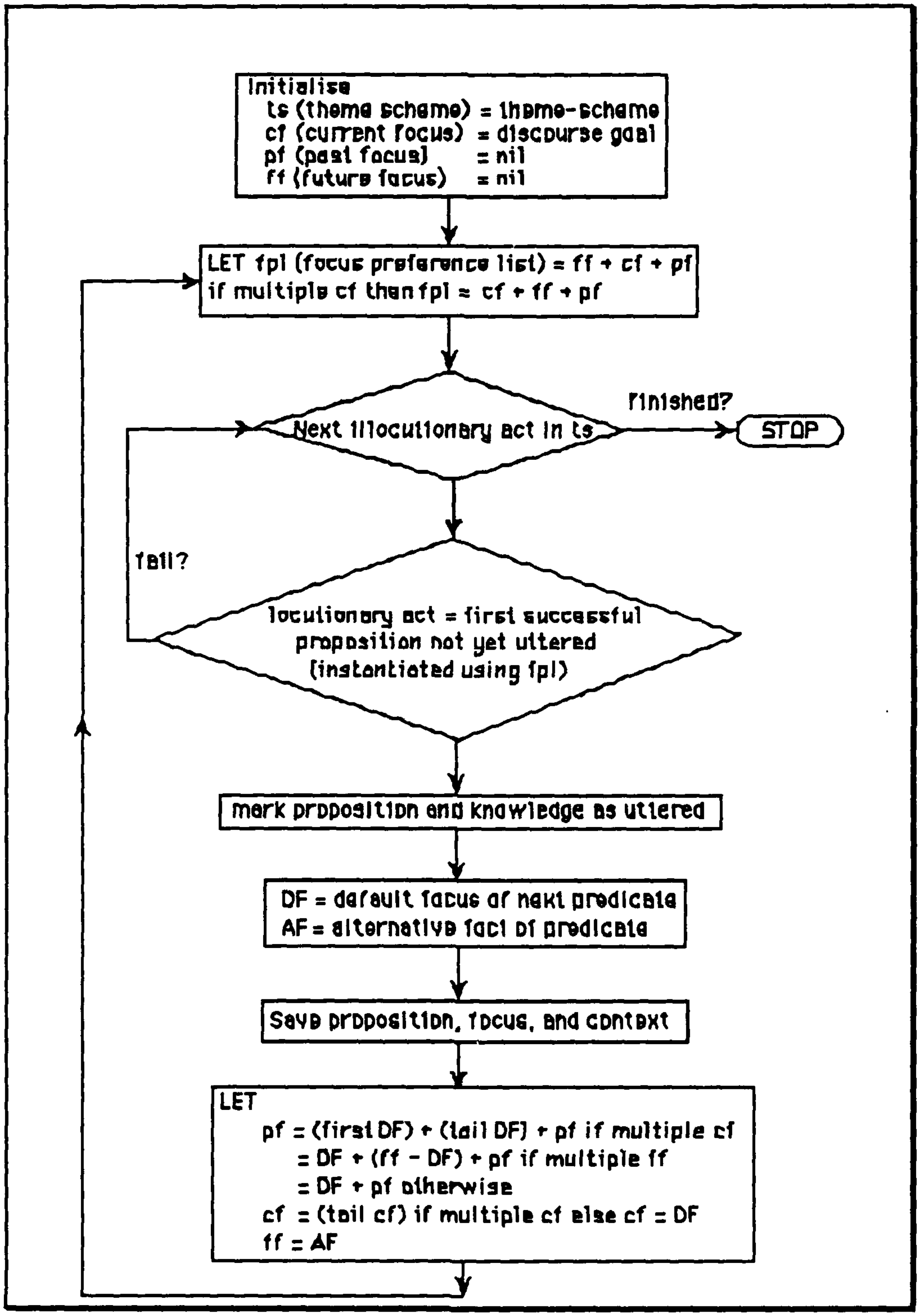

Figure 6.2 Predicate Selection Flow Chart. Selection constrained by focus.

6.4 Pragmatic Effects on Surface Form

Just as focus constraints augment text coherence, so grammatical choice constrained by context can act as a binder of discourse. In GENNY, context consists of *given* entities, mentioned previously in discourse, and *new* entities, introduced in the current utterance for the first time in discourse. Not rules but generalities govern the speaker's referential and grammatical choices [Brown and Yule, 1983, p. 189] with regard to content:

- speakers usually introduce *new* entities with indefinite referring expressions and with intonational prominence
- speakers usually refer to current *given* entities with attenuated syntactic and phonological forms

Exploiting these regularities in the contexts of discourse, hearers are able to interpret co-referential text. Conversely, these generalities allow us to make lexical decisions when building syntactic structures.

For example, when generating the first utterance in a define theme-scheme, where the discourse topic is brain, GENNY says *A brain is an organ located in the human skull.* Notice both the subject and object have indefinite articles as both are *new*. While it can be argued that the noun phrase within the prepositional phrase could also use an indefinite article, as it too represents new information, the adjective specifies a *human* skull and therefore the definite article is chosen.

Just as speakers utilize lexical devices to mark new information, so too *given* entities are referred to with attenuated syntactic forms. GENNY exploits given information to select definite noun phrases and anaphora (see section 8.3). For example, after introducing the entity, "brain", GENNY can refer to it as "the brain", since it is given. Furthermore, if "brain" is at the forefront of the intended hearer's mind (i.e. was the past CF), the anaphora can be used co-referring to it. This decision tacitly assumes the principle of analogy (things tend to be the same) together with the principle of local interpretation (change is minimal). Anaphor is discussed further in section 8.4.

Syntactically, focus suggests choice between active and passive constructs. There-insertion is used to promote the object to the subject position were the passive construction is not

possible (e.g. with a copula verb). It-extraposition can suggest focal stress (e.g. "It was John who hit Jill"). These are detailed in section 8.2. But first the rhetorical message must be interpreted by the semantic component, the first module of the threefold tactical generator which includes *semantics*, *grammatical relations*, and *syntax*.

Chapter 7

SEMANTICS

You do not understand this parable? How then are you going to understand other figures like it?
Mark 4:13,14

7.1 Introduction

Tactical generation components must map a rhetorical message onto surface form. In GENNY this process involves translation from the rhetorical proposition onto a *semantic case grammar* (this chapter), a *relational grammar* (chapter 8), a *syntactic grammar* (chapter 9), and finally onto surface form via morphology and orthography. Figure 7.1 relates these levels together with the previously discussed message formalism and pragmatics information. The motivation for these distinct levels of analysis is the lack of previous generators to map semantics onto syntax in a modular and well-motivated fashion (e.g. McKeown's hand-encoded dictionary of phrasal constituents).

7.2 Semantic Interpretation of Rhetorical Propositions

A variety of semantic representations are present in the literature including deep case relations, CD structures, and truth conditions or possible worlds [Fillmore, 1968; Schank and Abelson, 1977; Montague, 1974]. GENNY incorporates two of these meaning systems: Montague semantics for interpretation [Pulman, 1987] and case-based semantics [Fillmore, 1968, 1977] for generation. Montague semantics, implemented using the familiar λ-reduction

RHETORICAL PREDICATE

(definition ((left-hemisphere))
((region))
((location (brain human)) (function (feature-recognition))))

⇓

PRAGMATIC INFORMATION

PF	none
CF	left-hemisphere
FF	region
given	none

⇓

SEMANTIC INFORMATION

action	be
agent	left-hemisphere
patient	region
location	human brain
function	feature-recognition

⇓

RELATIONAL FUNCTION

predicate	be
subject	the left-hemisphere
object	a region
modifiers	for feature-recognition in the human brain

⇓

SYNTACTIC FUNCTION

the left-hemisphere
is
a region
for feature-recognition
located in the human brain

⇓

SURFACE FORM

The left-hemisphere is a region for feature-recognition located in the human brain.

Figure 7.1 Mapping from proposition to surface in GENNY.

mechanisms, rely on semantic entries for each lexical entry together with a semantic component for each grammatical rule. Conversely, the semantic role in the case representation is obtained from the rhetorical predicate.

GENNY translates the predicate into deep case roles of action, agent, patient, instrument, location, function, external location, beneficiary, manner, time, and state.[1] GENNY interprets the message formalism in three stages. First, the rhetorical predicate type is mapped onto an action guided by the function the utterance plays in discourse as well as the relationships of the entities in the message. Thus, a cause-effect message containing an object would utilize the verb have (e.g. *The brain has damage because* ...) whereas evidential knowledge would suggest other actions (e.g. *The instability observation is made because* ... or *The left-cognitive-flexibility symptom is manifest because* ...).

Secondly, case roles are selected based on their position in the message formalism. Finally, any modifiers which originate from the dda are interpreted using the semantic markers *location, external-location, function, instrument* which eventually translate to prepositional phrases of "located in", "on", "with", and "for". This treatment is certainly very limited, indeed testing revealed a need for more semantic markers and their corresponding deep case roles to represent and generate other surface forms (e.g. "from" for origin). This deep case semantics is fully documented in [Maybury, 1987b, Volume II].

The case formalism has received criticism that it is a mere notational variant of some preferred theory and at best is a mere taxonomy. [Fillmore, 1977, p. 70] clarifies the purpose of the deep case proposal as a recognition of a case-level organization of sentences rather than a complete grammatical model. He recognizes the need for "a level of representation including the grammatical relations subject and object." This level is represented as the relational function in GENNY which we now describe.

[1] A variety of case roles have been suggested (Fillmore 1968; Schank 1975; Grimes 1975). The case lists range in length from the most terse (nominative, ergative, locative) [Anderson, 1971], to a wider coverage illustrated recently [Sparck Jones and Boguraev, 1987].

Chapter 8

RELATIONAL FUNCTION

Man is but a network of relationships and these alone matter to him.
St. Exupéry

8.1 Introduction

Researchers in NL interpretation have recognized the utility of relational ideas. In GUS (Genial Understanding System) [Bobrow *et al.* 1977], for example, parsing is completed in two phases. First input is parsed into grammatical registers (subject, predicate, direct-object, indirect object) with prepositional phrases placed in a modifier list. Next the result is semantically interpreted using verb-case roles. Winograd [1983, p. 324] points out that in a language with a more developed case system (e.g. Russian and Japanese), the use of verb-centered analysis could be even more beneficial. Some inter-lingual studies also support a relational level of analysis [Perlmutter, 1980].

RG embodies a hierarchy of sentence participants so that in English, for example, the subject is 1, the direct-object is 2, and the indirect-object is 3. Rules can then capture generalities like: *to form the passive, promote 2 to 1 (direct-object to subject).* In this case the 1 element becomes chômeur (French for 'unemployed'), so it can either be dropped from the sentence, or transferred into a satellite phrase.

Current generators have largely ignored the promise of relational grammar. McKeown's dictionary component, for example, translat· knowledge base tokens into phrasal level

constituents via a hand-encoded dictionary [see McKeown, 1985, p. 167]. Clearly, this is linguistically insufficient, computationally expensive, and psychologically implausible. In contrast, GENNY has an independent representation of relational function, affording the power of relational grammar yet maintaining a well-tested, traditional phrase structure analysis.[1] GENNY employs syntactic experts to build grammatical components (e.g. subject, object, predicate) using both domain tokens and pragmatic information. For example, when forming noun phrases, indefinite articles are selected for new information whereas definite articles are preferred for given entities (discussed in section 8.3). (Even more sophisticated mechanisms are necessary to ensure use of minimal referring expressions while still uniquely identifying an object or concept in discourse.)

One obvious approach is to incorporate grammatical distinctions into syntactic grammars, for this certainly would decrease complexity. Within the standard transformational theory, for example, we could call the first noun phrase in a sentence its subject. However, this only marks the syntactic structures of the tree since 'subject' and 'indirect object' play no role at this level of representation. In more comprehensive paradigms (e.g. systemic or case[2] grammar), relational function plays a much greater role in the linguistic analysis.

8.2 Focal Stress and Surface Form

During generation, GENNY assigns focal prominence to relational constituents based on pragmatic constraints of relevancy. Assume, for example, a sentence is being generated where the message formalism translates to the semantic cases: subject ➜ alcoholism, object ➜ amnesia, and predicate ➜ causes. This might realize as *Alcoholism causes amnesia.*

Assume, however, that the focal shift algorithm determines that the next utterance is best described from the perspective of *amnesia*. The RG would indicate that to achieve this the 2

[1]Of course, one drawback of this approach is the computational expense of a full grammatical analysis. Accordingly, there is a speed versus completeness trade-off.

[2]Here case grammar, as opposed to deep case structure, describes a much wider range of grammatical phenomena: from deep to surface formats.

(object) should be promoted to 1 (subject). In the typical case, the predicate would be passivized (be + past participle of main verb), the preposition 'by' would be added before the new constituent of the 2 (object) register. Generation would eventually culminate in the surface form: *Amnesia is caused by alcoholism.*

However, there are some verbs (like the one in this sentence) which cannot be passivized. In these cases (e.g. be, have) syntactic ordering must account for focal prominence. So we can utter "It was a brain tumor that killed the patient (not a stroke)" to emphasize the semantic patient, "tumor". GENNY can utilize there-insertion and it-extraposition to achieve this type of forefronting.

Not only prominence (intonational or structural), but also lexical connectives can sew together discourse. The rhetorical function of an utterance in discourse suggests appropriate connectives (e.g. illustration ➡ "for example", cause-effect ➡ "because"). These are inserted at this relational level and serve not only as intrasentential markers, but more importantly, indicate the discourse role a sentence plays in the overall text.

8.3 Syntactic Experts

Relational constituents (subject, predicate, objects, and modifiers) are built with procedures which are experts in building syntactic phrases which realize these relational constituents. Provided the semantic message together with syntactic and pragmatic constraints, these procedures attempt to generate well-formed constituent phrases. Syntactic experts operate for three grammatical constituents in GENNY: noun phrases (NP), verb phrases (VP), and prepositional phrases (PP).

The NP builder, for example, consists of the pattern: **NP ➡ quantifier article adjective-list nominal-modifier-list head post-modifiers**. Articles are selected based on both syntactic constraints as well as pragmatic constraints of focus and context (given/new) as outlined in figure 8.1.

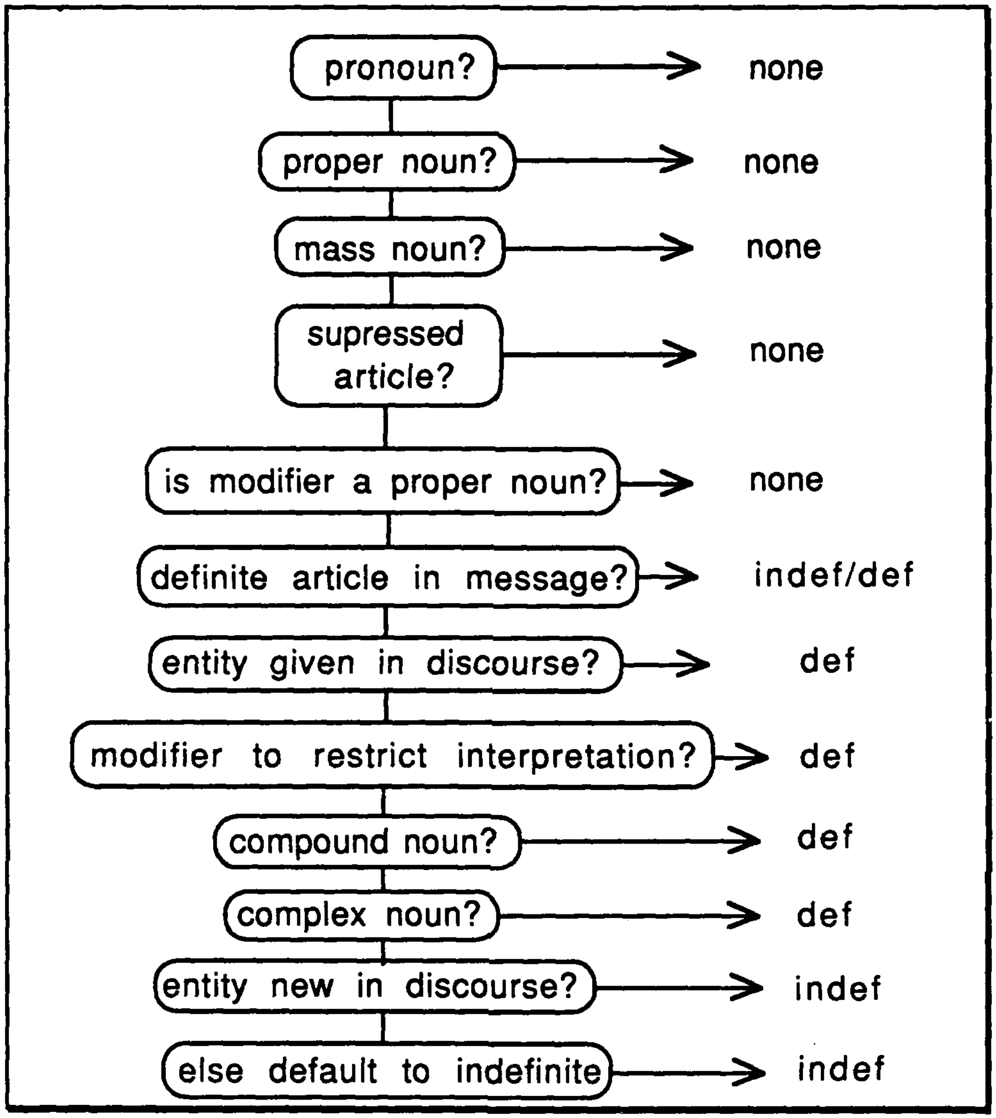

Figure 8.1 Article Selection Algorithm

For example, the syntactic specialist is able to generate the utterance *Vision is a symptom located in the left-occipital lobe with a function of recognizing images.* "Vision", a mass noun, requires no article. Also, GENNY's noun phrase specialist realizes that complex noun phrases composed of hyphenated words are distinguishable from simple nouns (e.g. *the left-occipital lobe* rather than *a left-occipital lobe*) (for complete algorithms see [Maybury, 1987b, section 9, volume II]). Note also that that the articles are morphologically consistent with the subsequent lexical item (discerning between "a" and"an"). It was found empirically that article agreement is dependent not just on the head of the noun phrase but on the subsequent linear word. This is language

dependent. In Italian, for example, articles agree with the head of the noun phrase but are modified by local morphology. Compare:

gli artisti storici	and	i bei negozi
the historic artists		the beautiful shops

This evidence supports the design of a modularized syntactic component interfaced to a language independent relational representation.

No examples of quantifiers were generated, although this would clearly be important in, for example, a logic knowledge formalism.[1] The adjective-list incorporates adjectives and ordinals while the nominal-modifier-list includes only nominals. Compound nouns were generated on the assumption that the message order passed from the semantic component would indicate the head noun as distinct from modifying nouns. This analysis was mirrored in the grammar (see grammar rule **np ➜ noun+noun** in [Maybury, 1987b, volume II, section 12.1]). The proper handling of compound nouns, however, is a major enterprise involving word sense, nominal phrase structure, and semantic word relations [Sparck-Jones, 1985].

The VP builder consists of the pattern **VP ➜ verb** or **VP ➜ auxiliary verb [past participle] particle**, depending upon the provided voice. An active voice will return the lexical entries for the provided semantic action. In contrast, a passive voice will indicate to the routine to select an appropriate auxiliary (e.g. "be") followed by the lexical entries for the verb,[2] followed by an appropriate particle if necessary, eventually to realize as "is contained in" or "is indicated by", for example.

Finally, a PP builder follows the pattern **PP ➜ preposition NP**, recursively calling the NP builder to complete its description (passing along pragmatic information). The preposition is provided to the routine by translating the semantic case role given with the entity. GENNY current

[1]Note, however, that quantifiers are computable from slot-filler type networks [McKeown, 1985].

[2]Lexical entries consist of only root or irregular forms of words. The feature list for the plural entry of the verb "contain", listed as (verb trans plur pres p3), is modified to (verb trans plur pres en) to form the past participle.

incorporates four case roles which eventually realize as PPs: location ("located in"), external-location ("on"), instrument ("with") and function ("for").

A nice property is that GENNY degrades gracefully when unable to translate or build certain phrasal constituents by attempting to utter what she can. For greater perspicuity, future work could investigate implementing GENNY's procedural syntactic experts as production rules as in [Tait, 1985]. Also, due to time constraints, work on lexical selection was limited [but see Sparck Jones and Tait, 1984].

8.4 Anaphora

The interpretation literature suggests that anaphor resolution should incorporate syntactic knowledge to constrain the search space and examine noun phrases in the immediately preceeding sentence first, followed by those of previous sentences [Hobbs, 1976 in Grishman, 1986]. Parse trees are searched breadth-first, top-down, and left-right so that subject and object are tested first. Recently, Sidner [1983] developed focus-based anaphor interpretation algorithms. Carter [1985], developed a shallow processing approach to anaphor resolution.

GENNY performs analysis for the restricted set of intersentential definite pronominal anaphora. It is in the NP builder that the decision to pronominalize is made. The algorithm to decide basically states that if the agent is in the list of past current foci and the agent is given, then pronominalize. Referring expressions are selected from set of possible pronominals by unifying syntactic features (person, number, gender, and animacy (proposed)).[1] During testing, GENNY produced:

```
A brain is an organ for understanding located in the human skull.
It has an importance value of ten.
It contains two regions: the left-hemisphere and the right-hemisphere.
The left-hemisphere, for example, has the feature-recognition function
located in it.
```

[1] See anaphora module in [Maybury, 1987b, Volume II, section 7] for details.

Of course the subject in sentence two and three is attenuated since it is forefronted in the reader's mind. It is interesting to note however, that the pronominalization in sentence four is ambiguous. In the message, "it" actually replaces "brain", yet in the utterance my own interpretation seems to favour resolution as "left-hemisphere". Apparently, longer texts will require reference mechanisms which incorporate more than just syntactic, recency and focus information. It seems that both locutionary as well as illocutionary context is necessary.[1]

In summary, RG serves as a natural symantico-syntax link. It promises to be a language independent representational level. In preliminary studies with Italian, RG appears a sufficiently robust formalism to handle at least simple active and passive Italian sentences within GENNY. Of course the lexicon, grammar, and syntactic specialists would have to be implemented for Italian, but the remaining (majority) of the system would remain constant. We now see how relational constituents are mapped onto surface form.

[1]If we introduce both "Alzheimer's disease" and "Huntington's disease" in discourse, subsequent nominal reference must uniquely identify the entity in discussion. The word "disease" is insufficient. Referential procedures are responsible for avoiding lexical ambiguity.

Chapter 9

SYNTACTIC FUNCTION

"When I use a word" Humpty Dumpty said, in a rather scornful tone,
"it means just what I choose it to mean -- neither more nor less."
"The question is," said Alice, "whether you *can* make words different things."
"The question is," said Humpty Dumpty, "which is to be master -- that's all."
Through the Looking Glass

9.1 Introduction

Within the functional paradigm there are two major approaches to generation at the syntactic level: systemic grammars and functional unification grammars. Systemic grammars [Halliday, 1976] distinguish between two levels of organization: *choice* and *structures* that realize choice. Language is classified as a network of systems and generation consists of selecting from alternatives.

One advantage of systemic grammars is efficiency. Unfortunately, there are several disadvantages. Systemic grammars introduce several complexities including lack of flexible ordering or omission, overlapping or discontinuous constituents, and agreement across systems (see [Winograd, 1983] for a detailed discussion). Even the largest systemic grammar does not have the breadth and clarity of current transformational grammars. It remains to be seen what systemic grammar will yield in grammatical coverage.

Conversely, unification grammars offer a well-tested formalism. Unfortunately, grammars of significant size are sluggish. Two alternatives are Functional Unification Grammar (FUG)

[Kay, 1979] and other non-functional Generalized Phrase Structure Grammar (GPSG) [Gazdar, 1982] which can encode function in feature-value pairs. While the former offers a uniform specification of function (semantic, grammatical, syntactic and lexical), it suffers many technical problems particularly with a grammar of any significant size. First, there are selection problems when alternatives are present[1] as well as problems with fragment generation. This remains an area for further exploration.

The alternative, GPSG, is well-studied and accounts for many complex phenomena including agreement and morphology, related forms, structural ambiguity, and unbounded movement. Meta-rules allow convenient description of generalities in rules. Features provide the possibility of including pragmatic registers (e.g. focus or given/new) directly in the grammar to allow the grammar to orchestrate a broader range of linguistic phenomena. Finally, semantic rules associated with individual grammar rules have shown promise in interpretation [Montague, 1974].

9.2 Grammar -- GPSG + Features

Phrase Structure Grammar (PSG) is based on an extension of Context Free Grammar (CFG). Typical rewrite rules such as "S ➡ NP VP" are augmented with features which constrain the possible well-formed syntactic trees. These rules can be sophisticated enough to cover agreement, morphology, missing/moved constituents, etc. For example, the active sentence level rule in GENNY is:

```
 S  [(type declarative) (voice active)]          ⇒
NP  [(count 1) (person 2) (gender 4)]            +
VP  [(count 1) (person 2) (tense 3) (voice active)]
```

For illustrative purposes, the capitalized characters indicate non-terminal symbols, followed by a list of feature-value pairs. Note that some feature values are symbols while others are variables

[1] McKeown who details FUG in TEXT, for example, side steps this problem by always taking the first successful alternative.

(integers) which indicate feature agreement. In the rule above, for example, the count (e.g. plural) and person (e.g. third-person) feature values must agree as indicated by variables 1 and 2. The voice feature would simply be changed to passive to state the top-level rule for passive sentences. The grammar includes rules for active and passive sentences, multi-sentential connectivity, and relative clauses, along with phrasal constructs (np, vp, pp, etc.). The documented grammar is listed in full in [Maybury, 1987b, volume II] along with mechanisms such as preparsers for efficiency.

For clarity, each rule has an associated name (**s<dec> ➡ np+vp**, for above). Also, each rule contains a λ-calculus meaning representation which is used to convert syntactic trees to logical form [Pulman, 1987]. This is intended for future interpretive use following the psycholinguistically motivated use of bi-directional grammars.[1]

9.3 Unification

The process of generation and (proposed) parsing is handled by the process of unification. Unification consists of using the grammar and features to build constituents which are placed on a well-formed sub-string table (WFSST) or chart [see Pulman, 1987 for detail]. The unifier percolates features up the chart (by matching and then binding feature variables), and generates all possible syntax trees from the given lexical entries. At the end of the generation, another routine simply reads off the completed trees (or partial trees, as in the case of ellipsis or fragments). The unbound variables in the syntax tree are bound with values from their agreeing constituents. The documented code for these routines can be found in Volume II, section 10.3.

9.4 Lexicon

The dictionary sub-system built for GENNY contains dictionary generation, access, edit, and removal functions. Lexical entries are listed in the format *<entry syntax semantics realization>*

[1] Some interesting work has been done using PROLOG with bi-directional grammars [Simons and Chester, 1982].

where *entry* refers to a token in the expert system, *syntax* includes categorical, agreement and morphological information, *semantics* includes a logical form meaning representation of the lexical item, and *realization* indicates the actual translation of the domain token into natural language. Variables were introduced into the syntax declarations to minimize repeat listing. Future plans include adding syntactic features of humanity, animacy, and abstractness for use in anaphor selection as well as in lexical selection (e.g. "who" or "which" insubordinate clauses).

To facilitate portability, a kernel dictionary was developed which contains frequently used words such as numbers, determiners, pronouns, prepositions, punctuation, conjunctions, connectives and core verbs. This was exploited when developing a second KB in photography for system evaluation.

9.5 Surface Morphology and Orthography

To complete the production, GENNY linearizes the output from the syntactic generator, synthesis lexical entries morphologically, then applies final orthographic conventions. Morphological synthesis is guided by syntactic features on lexical entries.

Orthographic conventions include text layout (spacing, pagination, new lines) and conventions such as capitalization and punctuation. Text layout was restricted to leaving a blank between lines. New lines were capitalized and punctuated. Use of pragmatic information at this level could suggest, for example, use of capitalization or exclamation marks for emphasis. Abbreviation also could be used for terseness when speaking to an expert.

9.6 Discussion

GPSG provides a clear and perspicuous syntactic formalism from which to implement syntax. While the current representation offers much promise, there are still many linguistic phenomena which require further investigation such as ellipsis, ill-formed language and structural ambiguity. Also, further interlingual investigations are necessary to fully realize the possibilities of

syntax independence. Finally, the problems involved in bi-directional grammars (e.g. lexicon development and semantic consistency) need to be closely examined.

Notwithstanding the need for extensive testing of these components, there appears to be both a theoretical and pragmatic bias toward this representation. The syntactic independence aids portability between languages. Moreover, the bi-direction of the grammar lends psychological credence with regard to cognitive efficiency. The scope and limitations of this formalism remain to be explored.

Chapter 10

TESTS AND EVALUATION

What is the difference between an optical-lens and an aperture?
An optical-lens is a component for focusing located in a camera.
It has a relative importance value of nine and a damage value of two.
An aperture is a component for light intensity control located in a lens.
It has a relative importance value of ten and a damage value of five.
An optical-lens and it have a different class, a similar type, and a different importance.
It and an aperture component, therefore, are similar entities.
GENNY August, 1987

10.1 Aim and Scope

The aim of GENNY was to produce connected and focused textual responses from a knowledge base in response to a simulated user request for information about or explanation of a topic. The scope for the project was limited to definitions, explanations and comparisons of KB entities.

10.2 Tests and Results

GENNY was tested by generating text for all three discourse goals (definition, explanation, comparison) for a variety of discourse topics (frames). Topics relating to frames were examined at all levels in the frame hierarchy. A second knowledge base and lexicon were developed to test claims of domain independency. Over fifty texts were generated from the system. A detailed trace of the system in operation and a representative output and are included in the Appendix.

GENNY generates well-focused and connected descriptions, explanations,and comparisons of objects within the provided knowledge base. The system failed to generated output (apologized) if the discourse goal was not represented or if the topic (frame) was not present in the knowledge base. Also, knowledge base token translation failed when lexical entries were not present in the dictionary, although the system degrades gracefully by attempting to realize what it was able to translate. The added distinguishing descriptive attributes had to be carefully hand-encoded or else errors would result in text (e.g. if the dda for brain was "(instrument understanding)" instead of "(function understanding)" we would get "The brain is a region with understanding" instead of "for understanding".

10.3 Evaluation

When asked to compare an optical-lens with an aperture, GENNY outputs the quote at the beginning of this section, which demonstrates results similar to that of McKeown's [1985] TEXT system (recognized as the state of the art in text generation and motivated by similar discourse needs). In response to a similar discourse goal as above, *What is the difference between a destroyer and a bomb?*, the TEXT system produces:

```
A destroyer is a surface ship with a DRAFT between  15 and 222.
A ship  is a vehicle.
A bomb is a free falling projectile that has a surface target location.
A free falling projectile is a lethal destructive device.
The bomb and the destroyer, therefore, are very different kinds of
entities.
```

GENNY produces produces similar definitions and comparisons as TEXT and, in addition, investigates explanations of knowledge base entities. This is partially a reflection of the richer (in terms of discourse goals) underlying application (expert systems versus data base systems). With a simulated request of *Why did you diagnose Korsakoff's disorder?*, GENNY responds:

```
Korsakoffs disorder is manifest because a memory-iq observation
and an apathetic observation indicate damage.
The memory-iq observation has a likelihood value of nine.
The apathetic observation has a likelihood value of ten.
```

Due to limited linguistic forms (lexical, sentential, and textual) GENNY's output can become boring. For example, the repetition of the attributive rhetorical predicate ("X has a damage value of five.") for all the constituent parts of an entity can lead to annoying textual replications. A greater number of possibilities in the schema should lead to richer and more varied text.

The claims of language independency were (minimally) tested by developing a small Italian dictionary, making minor modifications to the syntactic experts (e.g. position of adjectives in noun phrases), and modifying the morphological synthesizer. In response to the question, *What is a brain?*, GENNY uttered (English form in chapter 1):

```
Il cervello e una regione per comprensione situata nel cranio
umano. Il ha una valore di importanza relativa di dieci.
Il contiene due regioni: il emisphiro-della-sinistra e il
emisphiro-della-destra. Il emisphiro-della-sinistra ha una valore
di importanza relativa di dieci. Il emisphiro-della-destra ha una
valore di importanza relativa di dieci.
Il emisphiro-della-destra, per esempio, ha la funzione
comprensione-gestalt situata nel cervello destro.
```

While this output is grammatical and natural (as examined by a native Italian), the extent of GENNY's language independency requires rigorous testing.

10.4 Discussion

There are some linguistic phenomena handled by TEXT (e.g. quantification) which are not present in GENNY. This was a reflection of time constraints rather than a deficiency in the linguistic theory presented and could be incorporated in the future. GENNY is capable of generating the surface forms in McKeown's system (active, passive, and there-insertion sentences) but, in addition, it-extraposition for emphasis (driven by focus information).

GENNY includes mechanisms not present in TEXT, or for that matter in other NLG systems. GENNY refutes the fact that people always prefer future focus to current focus to past focus (FF > CF > PF) and instead prefers CF > FF > PF when there are multiple foci. Also, GENNY'S tactical component (as detailed in previous sections) is principled on a well-motivated translation from message formalism to surface form via relational grammar. In TEXT, no

linguistic analysis is performed on KB tokens: they are not translated but used directly in the text. Also, in GENNY, referring expressions (anaphor) and lexical choice (selection of indefinite and definite articles) was guided by context information (given/new).

Another difference lies in the representation of knowledge. McKeown had to hand encode both a generalization and attribute hierarchy. In contrast, KB modification for linguistic purposes in GENNY was modest (addition of a DDA for each entity). Investigation of a second KB (photography) demonstrated the domain independence of the system, offering support for the higher level linguistic theory.

Like TEXT, GENNY assumes a user input has been interpreted, and points into one or more frames in the knowledge base. Similarly, the discourse goal (e.g. explanation) is also assumed as in TEXT. Interpretation of input involves non-trivial issues of mapping the user query onto knowledge base entities and will have to be addressed in generation systems of the future.

The potential degree of system portability remains to be tested by interfacing to other applications such as a data base or a rule-based expert system. Furthermore, claims of language-independence must be fully tested. Extensive experimentation is still required to examine the robustness of the knowledge representation and knowledge selection procedures, particularly for expert systems outside of the causally related fault-diagnosis paradigm or those which have larger quantities of knowledge. Testing with even longer texts and contexts should reveal the efficacy of the text schema to impose a global framework and the local focus constraints to encourage local connectivity.

Chapter 11

CONCLUSION

Ancora imparo.
Michelangelo Buonarroti

11.1 Summary

This dissertation focuses on the key issue of NLG: generation under constraints. GENNY investigates these constraints on the spectrum from discourse to syntax. First a linguistically motivated framework of NLG was developed and then a computational model for realizing this was designed and implemented.

The linguistic issues investigated in GENNY include the analysis of common communicative strategies found in human-produced text and the well-motivated translation of a rhetorical message onto surface form. The computational model of generation implemented involved: the development and incorporation of high level text structures from natural texts; focus algorithms (global and local) for realization of the Gricean maxim of relevancy; a multi-level grammatical representation with particular emphasis on the role of language-independence; and mechanisms for improving textual coherence and plausibility (discourse plans, lexical connectives, and context-guided article selection).

11.2 Contributions

In contrast to previous work, GENNY incorporates both domain-independent linguistic structures (*theme-schemes* -- developed from analysis of natural texts) as well as a language-independent grammar formalism (RG). GENNY suggests algorithms for sticking to the point, moving from one focus to another, deciding what words to use, as well as deciding how to order them. GENNY also illustrates the promise of bi-directional grammars and dictionaries.

In theoretic terms, the system holds promise as a well-motivated linguistic representation which can be used for both generation and interpretation. In pragmatic terms, it is suggestive of a (domain and KR) portable and (language) universal system.

11.3 Limitations

The system's greatest limitation is the lack of and the limited pragmatic analysis (i.e. no user modeling, limited analysis of Gricean maxims). This both a reflection of time constraints coupled with a need for more theorctical research on these difficult higher level linguistic phenomena.

GENNY incorporates no creative expression. For example, old words could be coupled together to create new expressions utilizing the semantic lexical features together with some amalgamation routines. Also, there is no self-monitoring where the program "listens to itself" to detect ambiguity (lexical, structural, or referential). Furthermore, there is no post-editing for style to ensure a message or discourse realizes smoothly and cogently. Finally, the anaphoric analysis requires more sophisticated mechanisms which incorporate both locutionary and illocutionary knowledge. These issues suggest future paths of research.

11.4 Future Directions

The new frontiers include universality, discourse modelling (text coherence and cohesion), and audience modelling. Future generators need also to address pragmatic issues such as setting (e.g. speaker and hearer goals and relationships), and how they effect surface decisions. More sophisticated syntactic structures and their relation to focus need to be investigated including: parallel sentence structure, subordinate sentences, and textual connectives.

Only after the difficult issues of universality, discourse and pragmatics,and user modelling are fully tackled will effective practical generators emerge. Excellence in generation, however, awaits the synergism of formal knowledge with creativity. Then we will be able to translate indirect intention, deal sufficiently with co-reference, and generate not only connected and plausible,but also sophisticated text. But then again, Shakespeare didn't learn to write poetry overnight.

References

Allen, J., "Speech Synthesis from Unrestricted Text." In Fallside and Woods (eds.) *Computer Speech Processing*, 1985, pp. 461-477.

Allen J. (ed.), *From Text to Speech: The MITalk System*, Cambridge University Press, 1986.

Alshawi, H. "Memory and Context Mechanisms for Automatic Text Processing" (thesis), Report 60, Computer Laboratory, University of Cambridge, 1983. From lecture notes on *Language Applications*, Karen Sparck Jones, CSLP, CUED, Lent Term, 1987.

Anderson, J. M., *The Grammar of Case*, Cambridge University Press, 1971.

Appelt, D. *Planning English Sentences*, Cambridge University Press, 1985.

Bayes, Rev. T. An Essay Toward Solving a Problem in the Doctrine of Chance, *Philosophical Transactions of the Royal Society*, 1763.

de Beaugrande, R., *Text Production: Towards a Science of Composition*, Vol. XI in series **Advances in Discourse Processing**, Alex Publishing Corporation, 1984.

Becker, J. D. *The Phrasal Lexicon*, Bolt, Beranek and Newman, Technical report no. 3081, 1975.

Bobrow, D. and the PARC Understanding Group, "GUS, A Frame Driven Dialog System," *Artificial Intelligence 8*, 1977, 155-173 North-Holland. In Grosz, Sparck Jones and Webber, 1986.

Bossie, S. "A Tactical Component for Text Generation: Sentence Generation Using a Functional Grammar," University of Pennsylvania, Technical Report MS-CIS-81-5, Philadelphia, PA, 1981.

Bossie, S. and Mani, I. "An Overview of Research in Natural Language Generation," *TI Engineering Journal*, January-February 1986, pp. 52-57.

Brady M. and Berwick, R. (eds.) *Computational Models of Discourse*, Cambridge, Massachusetts: MIT Press, 1983.

Brown, G. and Yule, G., *Discourse Analysis*, Cambridge University Press, 1983.

Carter, D. M. "A Shallow Processing Approach to Anaphor Resolution, "University of Cambridge, Computer Laboratory, TR-88, 1985.

Chen, R. P. S., "The entity-relationship model -- towards a unified view of data," vol. 1, no. 1, 1976.

Cohen P., "On Knowing What to Say: Planning Speech Acts," Technical Report No. 118, University of Toronto, Toronto, 1978.

Cohen P., "The Need for Identification as a Planned Action," *Proceedings of the 7th Annual International Joint Conference on Artificial Intelligence*, 1981.

Cole, P. and Morgan, J. L (eds.), *Syntax and Semantics 3: Speech Acts*, Academic Press: New York, 1975.

Davey, A., *Discourse Production*, Edinburgh University Press, 1979.

van Dijk, T. A., *Text and Context*, London: Longman, 1977.

Dik, S. C., *Functional Grammar*, New York: North-Holland, 1978.

Dik, S. C., "Seventeen Sentences: Basic Principles and Applications of Functional Grammar." In *Syntax and Semantics 13*, Moravesik and Wirth (eds.), 1980, pp. 45-76.

Ehrlich, K. and Johnson-Laird, P. N., "Spatial Descriptions and Referential Continuity," *Journal of Verbal Learning and Verbal Behavior*, **21**, 1982, pp. 296-306.

Fallside, F. and Woods, W. A. (eds.) *Computer Speech Processing*, Prentice Hall, 1985.

Fillmore, C. J. "The Case for Case." In E. Bach and R. Harms (eds.) *Universals in Linguistic Theory*, New York: Holt, Rinehart and Winston, 1968.

Fillmore, C. J. "The Case for Case Reopened." *In Syntax and Semantics 18: Grammatical Relations,* P. Cole and J. Sadock (eds.), New York: Academic Press, 1977, pp. 59-81.

Fimbel, E., Gröscot, H, Lancel, J., Simonin, N., "Using a Text Model For Analysis and Generation," *Proceedings of the 2nd Annual Conference of ECACL*, 1985, pp. 226-231.

Gazdar, G., *Pragmatics: Implicature, Presupposition, and Logical Form*, New York: Academic Press, 1979.

Gazdar, G., "Phrase Structure Grammar." In P. Jacobson and G. K. Pullum (eds.) *On the Nature of Syntactic Representation*, Dordrecht: Reidel,1982.

Golden, C. J., "Computational Models of the Brain." In *Computers in Human Behavior*, Vol. 1, Pergamon Press, 1985, pp. 35-48.

Goldman, N. M., "Conceptual Generation," in Schank, R. C. (ed.), *Conceptual Information Process*, North-Holland: Amsterdam, 1975.

Granville, R., "Controlling Lexical Substitution in Computer Text Generation," *Proceedings of the 22nd Annual Conference of ACL*, 1984, pp. 381- 384.

Grice "Logic and Conversation." In Cole and Morgan, 1975, pp. 45-58.

Grimes, J. E., *The Thread of Discourse*, Mouton, The Hague, Paris,1975.

Grishman, R. and Hirschman, L., "Question-answering from Natural Language Medical Data Bases," *Artificial Intelligence 11*, 1978, pp. 25-43.

Grishman, R., "Response Generation in Question-answering Systems," *Proceedings of the 17th Annual Meeting of the ACL*, La Jolla, California, August, 1979, pp. 99-102.

Grishman, R., *Computational Linguistics: an Introduction,* Cambridge University Press, 1986.

Grosz, B. J., Sparck Jones, K. and Webber, B. L., *Readings in Natural Language Processing*, Los Altos, California: Morgan Kaufmann, 1986.

Grosz, B. J. "The Representation and Use of Focus in a System for Understanding Dialogs," *Proceedings of the Fifth Annual IJCAI*, Cambridge, Mass., pp. 67-76, Los Altos: William Kaufmann, 1977.

Halliday, M. A. K. *System and Function in Language*, London: Oxford University Press, 1976.

Halliday, M. A. K. *An Introduction to Functional Grammar*, London:Edward Arnold, 1985.

Halliday, M. A. K., and Hanan, R., *Cohesion in English*, London:Longman, 1976.

Hovy, E. H., "Generating Natural Language Under Pragmatic Constraints,"Ph. D. Dissertation, Yale University Department of Computer Science, March, 1987.

Jacobs, P., "PHRED: A Generator for Natural Language Interfaces," *Computational Linguistics*, Vol. 11, Number 4, October-December, 1985, pp. 219-242.

de Joia, A. and Stenton, A., *Terms in Linguistics: A Guide to Halliday*, London: Batsford Academic and Educational Ltd., 1980.

Johnson-Laird, P. N., *Mental Models*, Cambridge, Massachusetts: Harvard University Press, 1983.

Kant, E., *The Critique of Pure Reason*, second edition. Translated by J. M. D. Meiklejohn, London: Dent, 1934 (1787 original).

Kass, A. and Leake, D., "Types of Explanations," Yale University, CSD, Research Report #523, March, 1987.

Katz, J. J. *Propositional Structures and Illocutionary Force*, New York: Crowell, 1977.

Kay, M. "Functional Grammar." in *Proceedings of the 5th Annual Meeting of the Berkley Linguistic Society*, 1979.

Kukich, K., "Fluency in Natural Language Reports," 1984, to appear in *Readings in Natural Language Text Generation*, Bolc, L. (ed.),Springer-Verlag, 1986.

Kukich, K., "Feasibility of Automatic Natural Language Report Generation," *18th Annual Hawaii International Conference on Systems Sciences*, 2-4 January, 1985, U Hawaii, Honalulu, Hawaii.

Kukich, K., "Explanation Structures in XSEL," *Proceedings of the 23rd Meeting of the ACL*, Chicago, 1985.

Kukich, K., McDermott, J., and Wang, T., "Explaining XSEL's Reasoning," draft, 25 July 1985, Computer Science Department, Carnegie-Mellon University, Pittsburg, Pennsylvania.

Levinson, S., *Pragmatics*, Cambridge University Press, 1983.

Li, P. Y., Evens, M., Hier D., "Generating Medical Case Reports with the Linguistic String Parser," AAAI-86, *Proceedings of Fifth Annual Conference on AI*, Vol. II, Engineering, Philadelphia, Pennsylvania, August 11-15, 1986, pp. 1069-1073.

Mann, W. C. and Moore, J. A., "Computer Generation of Multi-paragraph English Text," *American Journal of Computational Linguistics*,Vol. 7, No. 1, January-March, 1981.

Mann, W. C., Bates, M., Grosz, B. J., McDonald, D. P, McKeown, K. R.,and Swartout, W. R. *Text Generation: The State of the Art and the Literature*, Technical Report 151/RR-81-101, Information Sciences Institute, Marina Del Ray, California, 1981.

Mann, W. C. "An Overview of the PENMAN text generation system," *Proceedings of the National Conference on Artificial Intelligence*,Washington, D. C., August, 1983, pp. 261-265.

Mathiesson, C. M. I. M., "A Grammar and Lexicon for a Text-production System," *Proceedings of the 19th Annual Meeting of the ACL*, Stanford, California, 1981, pp. 49-56.

Mauldin, M., "Semantic Rule Based Text Generation," *Proceedings of the 22nd Annual Meeting of the ACL*, 1984, pp. 376-380.

Maybury, M. T., "Artificial Intelligence: Generalized Expert Systems,"Fenwick Scholar Thesis, Department of Special Studies,College of the Holy Cross, Worcester, Massachusetts, May, 1986.

Maybury, M. T. "A Natural Language Interface to an Expert System for Neuropsychological Diagnosis," Report for Computer Speech and Language Processing Course, Cambridge University Engineering Department, February, 1987a.

Maybury, M. T. "A Report Generator," M Phil Dissertation, Cambridge University Engineering Department, September, 1987b.

Maybury, M. T. and Weiss, C. "The NEUROPSYCHOLOGIST Expert System Prototype," Conference Paper CIPS87-21, *Proceedings of the Canadian Information Processing Society*, Edmonton, Alberta, Canada, November 16-19, 1987c.

McCoy, K. "Correcting Object-Related Misconceptions." Ph D Thesis, University of Pennsylvania TR 85-57, December, 1985.

McDonald, D. D., "Language Production: The Source of the Dictionary," *Proceedings of the 19th Annual Conference of the ACL*, 1981, pp. 57-62.

McDonald, D. D., "Language Generation as a Computational Problem: an Introduction," COINS Technical Report 81-33, University of Massachusetts at Amherst, December, 1981.

McDonald, D. D. and Pustejovsky, J., "A Computational Theory of Prose Style for Natural Language Generation," *Proceedings of the Second Conference of European Chapter of ACL* 27-29 March,1985, pp. 187-193.

McDonald, D. D. and Pustejovsky, J., "TAG's as a Grammatical Formalism for Generation," *Proceedings of the 23rd Annual Meeting of the ACL*, 1985, pp. 94-103.

McDonald, D. D. and Pustejovsky, J., "Description-Directed Natural Language Generation," *Proceedings of the International Joint Conference on Artificial Intelligence*, 1985, pp. 799-805.

McDonald, D. D., "Natural Language Generation: Complexities and Techniques," in Nirenburg, S. (ed.), 1987. *Machine Translation:Theoretical and Methodological Issues*, from series **Studies in Natural Language Processing**, Cambridge University Press, 1987.

McDonald, D. D., "Natural Language as a Computational Problem: An Introduction." In Brady and Berwick, 1983.

McKeown, K. R., "Paraphrasing Using Given and New Information in a Question-Answering System," *Proceedings of the 17th Annual Meeting of ACL*, August, 1979, pp. 67-72.

McKeown, K. R., Wish, M., and Matthews, K., "Tailoring Explanations for the User," *Proceedings of the Annual International Joint Conference on Artificial Intelligence*, 1985, pp. 794-798.

McKeown, K. R., "Discourse Strategies for Generating Natural-Language Text," *Artificial Intelligence*, Elsevier Science Publishers,North-Holland, 1985, pp. 1-41.

McKeown, K. R., *Text Generation*, Cambridge University Press, 1985.

Meehan, J. R. "TALE-SPIN, an interactive Program that Writes Stories," in *Proceedings of the 5th Annual International Joint Conference on Artificial Intelligence*, August, 1977, pp. 91-98.

Minsky, M. "A Framework for Representing Knowledge," in P. H. Winston (ed.), *The Psychology of Computer Vision*, New York: McGraw-Hill, 1975.

Montague, J. L., *Formal Philosophy: Selected Papers*, (edited by R. H.Thomson), New Haven: Yale University Press, 1974.

Moravesik, E. A., and Wirth, J. R. (eds.) *Syntax and Semantics 13: Current Approaches to Syntax*, New York: Academic Press, 1980.

Nirenburg, S. (ed.), *Machine Translation: Theoretical and Methodological Issues,* from series **Studies in Natural Language Processing**, Cambridge University Press, 1987.

Parisi, D., "GEMS: A Model of Sentence Production," *Proceedings of the 2nd Annual Conference of ECACL*, 1985, pp. 258-262.

Perlmutter, D. "Relational Grammar." In *Syntax and Semantics 13*, Moravesik and Wirth (eds.), 1980, pp. 195-230.

Perlmutter, D. and Soames, S. *Syntactic Argumentation and the Structure of English*, Berkley: University of California Press, 1979.

Perlmutter, D. and Rosen, C. J. G. *Studies in Relational Grammar 2*, Chicago: University of Chicago Press, 1984.

Pulman, S. G., lecture notes for *Syntax and Parsing, Semantics and Inference*, and *Discourse Processing*, CSLP, CUED, Michaelmas and Lent Term, 1986-87.

Quirk, R. and Greenbaum, S., *A Concise Grammar of Contemporary English*, New York: Harcourt, Brace and Jovanovich, Inc., 1973.

Rosch, E., Classification of Real World Objects: Origins and Representations in Cognition. In P. N. Johnson-Laird and P. C. Wason (eds.) *Thinking:Readings in Cognitive Science*, Cambridge University Press, 1977.

Sanford, A. J. and Garrod, S.C., *Understanding Written Language*, Chichester:Wiley, 1981.

Searle, J. R., *Speech Acts*, Cambridge University Press, 1969.

Searle, J. R., *Indirect Speech Acts*. In Cole and Morgan, 1975, pp.59-82.

Schank, R. C., *Conceptual Information Processing*, New York: American Elsevier, 1975.

Schank, R. C. and Abelson, R. P., *Scripts, Plans, Goals, and Understanding*, Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.

Schank, R. C., *The Explanation Game*, Yale University CSD, Research Report #307, March, 1984.

Schank, R. C., *Explanation: A First Pass*, Yale University CSD, Research Report #330, September, 1984.

Schank, R. C. and Riesbeck, C., *Explanation: A Second Pass*, Yale University CSD, Research Report #384, July, 1985.

Scott, A. F., *Meaning and Style*, London: Macmillan, 1938.

Shortliffe, E. *Computer-based Medical Consultations: MYCIN*, Elsevier,1976.

Sidner, C. L. "Focusing in the Comprehension of Definite Anaphora." In Brady and Berwick, 1983, pp. 267-330.

Simmons, R. and Chester, D., "Relating Sentences and Semantic Networks with Procedural Logic," *Communications of the ACM* 25, 8, August, 1982, pp. 527-547.

Sparck Jones, K. "Compound Noun Interpretation Problems." In Fallside and Woods, *Computer Speech Processing*, 1985.

Sparck Jones, K. and Boguraev, B. K., "A Note on a Study of Cases," *Computational Linguistics*, Vol. 13, Numbers 1-2, January-June, 1987.

Sparck Jones, K. and Tait, J. I., "Linguistically Motivated Descriptive Term Selection," *Proceedings of COLING 84*, ACL, Stanford, 1984.

Swartout, W. R. "Producing Explanations and Justifications of Expert Consulting Programs," MIT Technical Report, MIT/LCS/TR-251, January, 1981.

Sykes, J. B. (ed.), *The Concise Oxford Dictionary of Current English*, seventh edition, Oxford University Press, 1984.

Tait, J. I., "An English Generator for a Case-Labeled Dependency Representation," *Proceedings of the 2nd Annual Conference of ECACL*, 1985, pp. 194-197.

Tait, J., "Generating Summaries Using a Script-Based Language Analyser."

The Varsity Handbook, 40th edition, Cambridge: Varsity Publications, 1986.

Walter, S. and Kalish C., "An Extensible Natural Language System,"ESD/MITRE Technical Objectives and Plans Report, Project 7590, Hanscome AFB, MA, October, 1985.

Weiner, J. L. "BLAH, a System Which Explains its Reasoning," *Artificial Intelligence 15*, 1980, pp. 19-48.

Williams, W. *Composition and Rhetoric*, D. C. Heath and Company:Boston, Massachusetts, 1893.

Winograd, T., *Understanding Natural Language*, Academic Press,Orlando, Florida, from **Cognitive Psychology 3**, No. 1, 1972.

Winograd, T., *Language as a Cognitive Process, Volume I: Syntax*, Addison-Wesley: Reading, Massachusetts, 1983.

Woods, W. A., "Transition Network Grammar for Natural Language Analysis," *Communications of the ACM*, Vol. 13,No. 10, October, 1970, pp. 591-606.

APPENDIX

System Traces and Output

System trace output is in regular typeface, user input is in **bold**, and descriptive commentary and generated text is in *italics*.

09:17:04 Tuesday, 1 September 1987

Franz Lisp, Opus 38.79-> (include main)
[fasl main.o]

The system first welcomes the user and describes the purpose of the GENNY.

Welcome to the GENNY text generation system for expert systems.
GENNY was designed to answer questions of the form:

-- What is an X?
-- Why did you diagnose Y? or Why does Y have a problem?
-- What is the difference between X and Y?

where X and Y are entities within the provided knowledge base.

These three types of questions are indicated by the keywords:
DEFINE, EXPLAIN, and COMPARE, respectively.

Next, the system asks the user to enter lexical and domain knowledge. These are the only two modules of domain specific knowledge which the generator employs. For example, the user could have replied **photography.dict** *and* **photography.kb** *if they wanted to interact with an expert photograpy fault diagnosis system.*

Please enter the domain dictionary file name? **neuropsychology.dict**
[load neuropsychology.dict]

What is the domain of discourse? **neuropsychology.kb**
[load neuropsychology.kb]

After domain specific knowledge is entered, the system asks for a top level discourse goal and a discourse topic entity. These reponses are the only input the generator requires to produce descriptions, comparisons, or explanations of objects in the domain.

Do you wish DEFINE, EXPLAIN, or COMPARE? **define**

What do you wish to know about? **brain**

Next, the system uses the top level goal (define, explain, or compare) to sketch out a very abstract plan of attack.

TEXT SKETCH:

introduction
description
example

Then the system culls out the pertinent information in the knowledge base by marking the topic (brain), its children (left-hemisphere and right-hemisphere), and its parent(s) (human) as explicitly in global focus. At this point the siblings of the discourse topic entity, brain, are marked implicitly in focus. The siblings include brother and sister nodes in the generalization hierarchy which in this example would include other organs found in the human body such as the heart and lungs.

SELECT KNOWLEDGE VISTA ==> ((brain) brain left-hemisphere right-hemisphere human)

GENERATE RELEVANT PROPOSITION POOL

Then the system uses the general text sketch produced above and, after reasoning about the discourse goal of the text and the available knowledge about the topic, it formulates a specific sequence of rhetorical relations which characterize the structure of the text. In our current example, the system decides upon the following sequence:

GENERATE DISCOURSE SKETCH:
(definition attributive constituent attributive attributive illustration)

Next the system uses the predicate selection algorithm (described in detail in Figure 6.2 in Chapter 6) to select predicates guided by the above sequence of rhetorical acts (called illocutionary acts in Figure 6.2).

GLOBAL FOCUS (TOPIC) ==> (brain)

```
LOCAL FOCUS CHOICES (FF/CF/PF) ==> (brain)
PREDICATE SELECTED ==>
(definition  ((brain))
             ((region))
             ((location (skull human)) (function (understanding))))

LOCAL FOCUS CHOICES (FF/CF/PF) ==> (region brain)
PREDICATE SELECTED ==>
(attributive  ((brain))
              ((value importance indef ten relative)))

LOCAL FOCUS CHOICES (FF/CF/PF) ==> (value brain)
PREDICATE SELECTED ==>
(constituent  ((brain))
              ((region two none))
              nil
              ((region left-hemisphere) (region right-hemisphere)))

LOCAL FOCUS CHOICES (FF/CF/PF) ==> (region left-hemisphere right-hemisphere brain)
PREDICATE SELECTED ==>
(attributive  ((left-hemisphere))
              ((value importance indef ten relative)))

LOCAL FOCUS CHOICES (FF/CF/PF) ==> (value left-hemisphere region right-hemisphere
brain)
PREDICATE SELECTED ==>
(attributive  ((right-hemisphere))
              ((value importance indef ten relative)))
```

LOCAL FOCUS CHOICES (FF/CF/PF) ==> (value right-hemisphere region left-hemisphere brain)
PREDICATE SELECTED ==>
(illustration ((region right-hemisphere))
((function gestalt-understanding)))

After the system selects the appropriate rhetorical acts, each of which are filled with information from the knowledge base, the system then recursively translates them onto surface form:

```
= = = = = = = = = = = = = = = = = = = = = = = = = = =
= = = = = = = = RHETORICAL PREDICATE = = = = = = =
= = = = = = = = = = = = = = = = = = = = = = = = = = =
```

(definition ((brain))
((region))
((location (skull human)) (function (understanding))))

PRAGMATIC FUNCTION (discourse-topic-entity/focus/given) :

((brain) (nil (brain) (region)) nil)

SEMANTIC FUNCTION:
action agent patient inst loc funct manner time
(be ((brain)) ((region)) nil (skull human) (understanding) nil nil nil nil)

RELATIONAL FUNCTION (voice and form): (active)

LEXICAL INPUT TO SENTENCE GENERATOR:

((a ((determiner count sing3p indefart notof noneg nonum) (article before consonant) a))

(brain ((noun count sing3p neuter) region brain))

(be ((copula plur pres p3) (L (_P) (L (_WH) (_P (L (_y) (equal _WH _y))))) are) ((copula sing3p pres p3) (L (_P) (L (_WH) (_P (L (_y) (equal _WH _y))))) is) ((copula sing pres p1) (L (_P) (L (_WH) (_P (L (_y) (equal _WH _y))))) am))

(a ((determiner count sing3p indefart notof noneg nonum) (article before consonant) a))

(region ((noun count 1 neuter) region region))

(for ((connective for-example) for for) ((preposition) (indicating purpose) for))

(understanding ((noun mass 1 neuter) consciousness understanding))

(located ((preposition located-in) (located-in) located))

(in ((preposition en) (contained-in) in) ((preposition located-in) (located-in) in) ((preposition) (inner or inward location) in))

(the ((determiner count 1 defart notof noneg nonum) (sing/plur form of the) the))

(human ((noun count 1 neuter) human human))

(skull ((noun count 1 neuter) (cranial container and protector) skull)))

SYNTAX OUTPUT FROM SENTENCE GENERATOR:

(((s declarative active)
((np sing3p p3 neuter) ((determiner count sing3p indefart notof noneg nonum) ((a))) ((n1 sing3p neuter) ((noun count sing3p neuter) ((brain)))))
((vp sing3p p3 pres active) ((copula sing3p pres p3) ((is))) ((np sing3p p3 neuter) ((np sing3p p3 neuter) ((np sing3p p3 neuter) ((determiner count sing3p indefart notof noneg nonum) ((a))) ((n1 sing3p neuter) ((noun count sing3p neuter) ((region))))) ((pp) ((preposition) ((for))) ((np sing3p p3 neuter) ((noun mass sing3p neuter) ((understanding)))))) ((pp) ((preposition located-in) ((located))) ((preposition located-in) ((in))) ((np 27 p3 neuter) ((determiner count 15 defart notof noneg nonum) ((the))) ((n1 27 neuter) ((noun count 21 neuter) ((human))) ((noun count 27 neuter) ((skull)))))))))

((s declarative active)
((np sing3p p3 neuter) ((determiner count sing3p indefart notof noneg nonum) ((a))) ((n1 sing3p neuter) ((noun count sing3p neuter) ((brain)))))
((vp sing3p p3 pres active) ((copula sing3p pres p3) ((is))) ((np sing3p p3 neuter) ((np sing3p p3 neuter) ((determiner count sing3p indefart notof noneg nonum) ((a))) ((n1 sing3p neuter) ((noun count sing3p neuter) ((region))))) ((pp) ((preposition) ((for))) ((np sing3p p3 neuter) ((np sing3p p3 neuter) ((noun mass sing3p neuter) ((understanding)))) ((pp) ((preposition located-in) ((located))) ((preposition located-in) ((in))) ((np 27 p3 neuter) ((determiner count 15 defart notof noneg nonum) ((the))) ((n1 27 neuter) ((noun count 21 neuter) ((human))) ((noun count 27 neuter) ((skull)))))))))))))t

= =
= = = = = = = = RHETORICAL PREDICATE = = = = = = = =
= =|
(attributive ((brain)) ((value importance indef ten relative)))

PRAGMATIC FUNCTION (discourse-topic-entity/focus/given) :

((brain) (((brain)) (brain) (value)) (brain region))

SEMANTIC FUNCTION:
action agent patient inst loc funct manner time
(have ((brain)) ((value importance indef ten relative)) nil nil nil nil nil nil nil)

RELATIONAL FUNCTION (voice and form) : (active)

LEXICAL INPUT TO SENTENCE GENERATOR:

((it ((pronoun pers sing3p subj p3 neuter) (a thing) it))

(have ((have-v sing3p pres p3) (to own or posess - irregular |3p| sing) has)
((have-v plur pres p1) (to own or posess) have)
((have-v sing pres p1) (to own or posess) have))

(a ((determiner count sing3p indefart notof noneg nonum) (article before consonant) a))

(relative ((adjective attributive) relative relative))

(importance ((noun count 1 neuter) importance importance))

(value ((noun count 1 neuter) value value))

(of ((preposition) (place of origin) of))

(ten ((number plur) (lexical representation of number 10) ten)))

SYNTAX OUTPUT FROM SENTENCE GENERATOR:

(((s declarative active)
((np sing3p p3 neuter) ((pronoun pers sing3p subj p3 neuter) ((it))))
((vp sing3p p3 pres active) ((have-v sing3p pres p3) ((has))) ((np sing3p p3 neuter) ((np sing3p p3 neuter) ((determiner count sing3p indefart notof noneg nonum) ((a))) ((n1 sing3p neuter) ((adjp attributive) ((adjective attributive) ((relative)))) ((n1 sing3p neuter) ((noun count 3 neuter) ((importance))) ((noun count sing3p neuter) ((value)))))) ((pp) ((preposition) ((of))) ((number plur) ((ten))))))))t

= =
= = = = = = = = RHETORICAL PREDICATE = = = = = = =
= =

(constituent ((brain))
((region two none))
nil
((region left-hemisphere) (region right-hemisphere)))

PRAGMATIC FUNCTION (discourse-topic-entity/focus/given) :

((brain) (((brain) (brain)) (brain) (region left-hemisphere right-hemisphere)) (brain value region))

SEMANTIC FUNCTION:
action agent patient inst loc funct manner time
(contain ((brain)) ((region two none)) nil nil nil nil ((region left-hemisphere) (region right-hemisphere)) nil nil)

RELATIONAL FUNCTION (voice and form): (active colon-insertion)

LEXICAL INPUT TO SENTENCE GENERATOR:

((it ((pronoun pers sing3p subj p3 neuter) (a thing) it))

(contain ((trans sing3p pres p3) (restricted or otherwise limited) contain)
((trans plur pres p3) (restricted or otherwise limited) contain))

(two ((number plur) (lexical representation of number 2) two))

(region ((noun count 1 neuter) region region)) (colon ((colon) colon colon))

(the ((determiner count 1 defart notof noneg nonum) (sing/plur form of the) the))

(left-hemisphere ((noun count sing3p neuter) region left-hemisphere))

(region ((noun count 1 neuter) region region))

(and ((conjunction coord) (intersection) and))

(the ((determiner count 1 defart notof noneg nonum) (sing/plur form of the) the))

(right-hemisphere ((noun count sing3p neuter) region right-hemisphere))

(region ((noun count 1 neuter) region region)))

SYNTAX OUTPUT FROM SENTENCE GENERATOR:

(((s declarative active)
((np sing3p p3 neuter) ((pronoun pers sing3p subj p3 neuter) ((it))))
((vp sing3p p3 pres active) ((trans sing3p pres p3) ((contain))) ((np plur p3 neuter) ((np plur p3 neuter) ((number plur) ((two))) ((n1 plur neuter) ((noun count plur neuter) ((region))))) ((colon) ((colon))) ((np plur p3 neuter) ((np 15 p3 neuter) ((determiner count 9 defart notof noneg nonum) ((the))) ((n1 15 neuter) ((noun count sing3p neuter) ((left-hemisphere))) ((noun count 15 neuter) ((region))))) ((conjunction coord) ((and))) ((np 27 p3 neuter) ((determiner count 21 defart notof noneg nonum) ((the))) ((n1 27 neuter) ((noun count sing3p neuter) ((right-hemisphere))) ((noun count 27 neuter) ((region)))))))))

((s declarative active)
((np sing3p p3 neuter) ((pronoun pers sing3p subj p3 neuter) ((it))))
((vp sing3p p3 pres active) ((trans sing3p pres p3) ((contain))) ((np plur p3 neuter) ((np plur p3 neuter) ((np plur p3 neuter) ((number plur) ((two))) ((n1 plur neuter) ((noun count plur neuter) ((region))))) ((colon) ((colon))) ((np 15 p3 neuter) ((determiner count 9 defart notof noneg nonum) ((the))) ((n1 15 neuter) ((noun count sing3p neuter) ((left-hemisphere))) ((noun count 15 neuter) ((region)))))) ((conjunction coord) ((and))) ((np 27 p3 neuter) ((determiner count 21 defart notof noneg nonum) ((the))) ((n1 27 neuter) ((noun count sing3p neuter) ((right-hemisphere))) ((noun count 27 neuter) ((region)))))))))t

= =
= = = = = = = = RHETORICAL PREDICATE = = = = = = =
= =

(attributive ((left-hemisphere)) ((value importance indef ten relative)))

PRAGMATIC FUNCTION (discourse-topic-entity/focus/given) :

((left-hemisphere) (((brain) (brain) (brain)) (left-hemisphere) (value)) (brain region left-hemisphere right-hemisphere value))

SEMANTIC FUNCTION:
action agent patient inst loc funct manner time
(have ((left-hemisphere)) ((value importance indef ten relative)) nil nil nil nil nil nil nil)

RELATIONAL FUNCTION (voice and form): (active)

LEXICAL INPUT TO SENTENCE GENERATOR:

((the ((determiner count 1 defart notof noneg nonum) (sing/plur form of the) the))

(left-hemisphere ((noun count sing3p neuter) region left-hemisphere))

(have ((have-v sing3p pres p3) (to own or posess - irregular |3p| sing) has)
((have-v plur pres p1) (to own or posess) have)
((have-v sing pres p1) (to own or posess) have))

(a ((determiner count sing3p indefart notof noneg nonum) (article before consonant) a))

(relative ((adjective attributive) relative relative))

(importance ((noun count 1 neuter) importance importance)) (value ((noun count 1 neuter) value value))

(of ((preposition) (place of origin) of))

(ten ((number plur) (lexical representation of number 10) ten)))

SYNTAX OUTPUT FROM SENTENCE GENERATOR:

(((s declarative active)
((np sing3p p3 neuter) ((determiner count 3 defart notof noneg nonum) ((the))) ((n1 sing3p neuter) ((noun count sing3p neuter) ((left-hemisphere)))))
((vp sing3p p3 pres active) ((have-v sing3p pres p3) ((has))) ((np sing3p p3 neuter) ((np sing3p p3 neuter) ((determiner count sing3p indefart notof noneg nonum) ((a))) ((n1 sing3p neuter) ((adjp attributive) ((adjective attributive) ((relative)))) ((n1 sing3p neuter) ((noun count 9 neuter) ((importance))) ((noun count sing3p neuter) ((value)))))) ((pp) ((preposition) ((of))) ((number plur) ((ten))))))))t

= =
= = = = = = = = RHETORICAL PREDICATE = = = = = = =
= =
(attributive ((right-hemisphere)) ((value importance indef ten relative)))

PRAGMATIC FUNCTION (discourse-topic-entity/focus/given) :

((right-hemisphere) (((region left-hemisphere right-hemisphere) (brain) (brain) (brain)) (right-hemisphere) (value)) (left-hemisphere value brain region right-hemisphere))

SEMANTIC FUNCTION:
action agent patient inst loc funct manner time
(have ((right-hemisphere)) ((value importance indef ten relative)) nil nil nil nil nil nil nil)

RELATIONAL FUNCTION (voice and form) : (active)

LEXICAL INPUT TO SENTENCE GENERATOR:

((the ((determiner count 1 defart notof noneg nonum) (sing/plur form of the) the))

(right-hemisphere ((noun count sing3p neuter) region right-hemisphere))

(have ((have-v sing3p pres p3) (to own or posess - irregular l3pl sing) has)
((have-v plur pres p1) (to own or posess) have)
((have-v sing pres p1) (to own or posess) have))

(a ((determiner count sing3p indefart notof noneg nonum) (article before consonant) a))

(relative ((adjective attributive) relative relative))

(importance ((noun count 1 neuter) importance importance))

(value ((noun count 1 neuter) value value))

(of ((preposition) (place of origin) of))

(ten ((number plur) (lexical representation of number 10) ten)))

SYNTAX OUTPUT FROM SENTENCE GENERATOR:

(((s declarative active)
((np sing3p p3 neuter) ((determiner count 3 defart notof noneg nonum) ((the))) ((n1 sing3p neuter) ((noun count sing3p neuter) ((right-hemisphere)))))
((vp sing3p p3 pres active) ((have-v sing3p pres p3) ((has))) ((np sing3p p3 neuter) ((np sing3p p3 neuter) ((determiner count sing3p indefart notof noneg nonum) ((a))) ((n1 sing3p neuter) ((adjp attributive) ((adjective attributive) ((relative)))) ((n1 sing3p neuter) ((noun count 9 neuter) ((importance))) ((noun count sing3p neuter) ((value)))))) ((pp) ((preposition) ((of))) ((number plur) ((ten))))))))t

= =
= = = = = = = = RHETORICAL PREDICATE = = = = = = =
= =

(illustration ((region right-hemisphere)) ((function gestalt-understanding)))

PRAGMATIC FUNCTION (discourse-topic-entity/focus/given) :

((right-hemisphere) (((right-hemisphere) (region left-hemisphere right-hemisphere) (brain) (brain) (brain)) (right-hemisphere) (gestalt-understanding)) (right-hemisphere value left-hemisphere brain region))

SEMANTIC FUNCTION:
action agent patient inst loc funct manner time
(have ((region right-hemisphere)) ((function gestalt-understanding)) nil nil nil nil nil nil)

RELATIONAL FUNCTION (voice and form) : (active example-insertion)

LEXICAL INPUT TO SENTENCE GENERATOR:

((the ((determiner count 1 defart notof noneg nonum) (sing/plur form of the) the))

(right-hemisphere ((noun count sing3p neuter) region right-hemisphere))

(region ((noun count 1 neuter) region region))

(comma ((comma) comma comma))

(for ((connective for-example) for for) ((preposition) (indicating purpose) for))

(example ((connective for-example) example example)
((noun mass 1 neuter) example example))

(comma ((comma) comma comma))

(have ((have-v sing3p pres p3) (to own or posess - irregular |3p| sing) has)
((have-v plur pres p1) (to own or posess) have)
((have-v sing pres p1) (to own or posess) have))

(the ((determiner count 1 defart notof noneg nonum) (sing/plur form of the) the))

(gestalt-understanding ((noun mass 1 neuter) function gestalt-understanding))

(function ((noun count 1 neuter) function function)
((verb trans sing3p pres p1) (telling) function)
((verb trans plur pres p3) (telling) function)))

SYNTAX OUTPUT FROM SENTENCE GENERATOR:

(((s declarative active) ((np 9 p3 neuter)
((np 9 p3 neuter) ((determiner count 3 defart notof noneg nonum) ((the))) ((n1 9 neuter) ((noun count sing3p neuter) ((right-hemisphere))) ((noun count 9 neuter) ((region))))) ((comma) ((comma))) ((rel for-example) ((connective for-example) ((for))) ((connective for-example) ((example)))) ((comma) ((comma))))
((vp sing3p p3 pres active) ((have-v sing3p pres p3) ((has))) ((np 33 p3 neuter) ((determiner count 21 defart notof noneg nonum) ((the))) ((n1 33 neuter) ((noun mass 27 neuter) ((gestalt-understanding))) ((noun count 33 neuter) ((function)))))))

((s declarative active)
((np 9 p3 neuter) ((np 9 p3 neuter) ((determiner count 3 defart notof noneg nonum) ((the))) ((n1 9 neuter) ((noun count sing3p neuter) ((right-hemisphere))) ((noun count 9 neuter) ((region))))) ((comma) ((comma))) ((rel for-example) ((connective for-example) ((for))) ((connective for-example) ((example)))) ((comma) ((comma))))
((vp plur p3 pres active) ((have-v plur pres p1) ((have))) ((np 33 p3 neuter) ((determiner count 21 defart notof noneg nonum) ((the))) ((n1 33 neuter) ((noun mass 27 neuter) ((gestalt-understanding))) ((noun count 33 neuter) ((function)))))))

((s declarative active)
((np 9 p3 neuter) ((np 9 p3 neuter) ((determiner count 3 defart notof noneg nonum) ((the))) ((n1 9 neuter) ((noun count sing3p neuter) ((right-hemisphere))) ((noun count 9 neuter) ((region))))) ((comma) ((comma))) ((rel for-example) ((connective for-example) ((for))) ((connective for-example) ((example)))) ((comma) ((comma))))
((vp sing p3 pres active) ((have-v sing pres p1) ((have))) ((np 33 p3 neuter) ((determiner count 21 defart notof noneg nonum) ((the))) ((n1 33 neuter) ((noun mass 27 neuter) ((gestalt-understanding))) ((noun count 33 neuter) ((function))))))))t

DISCOURSE STRUCTURE + FOCUS + GIVEN

(((definition ((brain)) ((region)) ((location (skull human)) (function (understanding)))) (nil (brain) (region)) nil)

((attributive ((brain)) ((value importance indef ten relative))) (((brain)) (brain) (value)) (brain region))

((constituent ((brain)) ((region two none)) nil ((region left-hemisphere) (region right-hemisphere))) (((brain) (brain)) (brain) (region left-hemisphere right-hemisphere)) (brain value region))

((attributive ((left-hemisphere)) ((value importance indef ten relative))) (((brain) (brain) (brain)) (left-hemisphere) (value)) (brain region left-hemisphere right-hemisphere value))

((attributive ((right-hemisphere)) ((value importance indef ten relative))) (((region left-hemisphere right-hemisphere) (brain) (brain) (brain)) (right-hemisphere) (value)) (left-hemisphere value brain region right-hemisphere))

((illustration ((region right-hemisphere)) ((function gestalt-understanding))) (((right-hemisphere) (region left-hemisphere right-hemisphere) (brain) (brain) (brain)) (right-hemisphere) (gestalt-understanding)) (right-hemisphere value left-hemisphere brain region)))

After semantic interpretation of the rhetorical messages, relational grammar analysis, and syntactic tree generation, the system then reads syntactic entries off of the tree depth-first and uses morphological synthesis routines to produce final lexemes.

MESSAGE REALIZATION

((a brain is a region for understanding located in the human skull)
(it has a relative importance value of ten)
(it contains two regions colon the left-hemisphere region and the right-hemisphere region)
(the left-hemisphere has a relative importance value of ten)
(the right-hemisphere has a relative importance value of ten)
(the right-hemisphere region comma for example comma has the gestalt-understanding function))

These lexemes are the then formatted by orthographic routines which add appropriate punctuation and spacing to produce the final surface form:

SURFACE FORM

A brain is a region for understanding located in the human skull.
It has a relative importance value of ten.
It contains two regions: the left-hemisphere region and the right-hemisphere region.
The left-hemisphere has a relative importance value of ten.
The right-hemisphere has a relative importance value of ten.
The right-hemisphere region, for example, has the gestalt-understanding function.

-> (exit)

Index

MISSION
of
Rome Air Development Center

RADC plans and executes research, development, test and selected acquisition programs in support of Command, Control, Communications and Intelligence (C^3I) activities. Technical and engineering support within areas of competence is provided to ESD Program Offices (POs) and other ESD elements to perform effective acquisition of C^3I systems. The areas of technical competence include communications, command and control, battle management information processing, surveillance sensors, intelligence data collection and handling, solid state sciences, electromagnetics, and propagation, and electronic reliability/maintainability and compatibility.